Mathematica®
Progress Together

Independent Evaluation of Comprehensive Primary Care Plus (CPC+):

Appendices to the Final Report, Volume II

December 2023

Independent Evaluation of Comprehensive Primary Care Plus (CPC+):

Appendices to the Final Report, Volume II

December 2023

Lead authors:

Jessica Laird, Lori Timmins, Ning Fu, Julia Embry, Brianna Sullivan, Priya Shanmugam, Nikkilyn Morrison, Geraldine Haile, Eugene Rich, Caroline Mack, Ellie Holland, Arkadipta Ghosh, Melanie Au, Breanna Arndt, Jelena Zurovac, Irina Degtiar, Eunhae Shin, Lianlian Lei, Amanda Markovitz, Carol Urato, Wenjia Zhu, Anthony Austin, Joel Earlywine, Christopher Fleming, Mariel Finucane, Priya Bajaj, Sean Orzol, Stacy Dale, Pragya Singh, Rosalind Keith, Ha Tu, Nancy Duda, Kristin Geonnotti, Mynti Hossain, Genna Cohen, Nancy McCall, Ann O'Malley, Dana Petersen

Contributing authors (in alphabetical order):

Rocky Aikens, Mattan Alalouf, Sophia Artis, Megan Baker, Jennifer Baskwell, Tom Bell, Laura Blue, Ellen Bouchery, Jonathan Brown, Sarah Brunskill, Claire Burkhart, Saumya Chatrath, Christine Cheu, Jared Coopersmith, Eric Dehus, Kathryn Dewitt, James Drury, Sheena Flowers, Jan Genevro, Cindy George, Gillian Giudice, Sabitha Gopalsamy, Laura Hanson, Natalie Hazelwood, Heather Ingraham, Alex Kobrin, Timothy Lake, Mark Lee, Marlena Luhr, Jeff McCartney, Andrew McGuirk, Katie Morrison Lee, Keep Nathanson, Maya Palakal, Christopher Palo, Margaret Raskob, Carol Razafindrakoto, Lauryn Ringwood, Michelle Roozeboom-Baker, Tyler Rose, Mike Rudacille, Rumin Sarwar, John Schurrer, Christina Sintek, Xiaofan Sun, Kaylyn Swankoski, Tobi Tandeciarz, Dan Thal, Jacob Thomas, Haixia Xu

Submitted to:

U.S. Department of Health and Human Services

Center for Medicare & Medicaid Innovation

7500 Security Blvd.

Baltimore, MD 21244-1850

Project Officer: Timothy Day

Contract number: HHSM-500-2014-00034I/ HHSM-500-T0010

Submitted by:

Mathematica

P.O. Box 2393

Princeton, NJ 08543-2393

Telephone: (609) 799-3535

Facsimile: (609) 799-0005

Project Director and Principal Investigator: Nancy McCall

Deputy Project Director: Nancy Duda

Co-Principal Investigators: Ann O'Malley and Stacy Dale

Reference Number: 50319

Contents

Tables

5.A. Detailed impact results over the five years of CPC+

5.A.1. Medicare FFS service use

5.A.2. Medicare FFS expenditures

Figures

Chapter 5 Appendix

5.A. Detailed impact results over the five years of CPC+

This Appendix supplements the main chapter by providing yearly impact estimates as well as detailed findings from subgroup analyses, sensitivity tests, and aggregate impact results for expenditures, service use, and quality of care outcomes. We focus on those practices that started in 2017.[1] For each set of outcomes, we report all the detailed findings for Track 1 CPC+ practices followed by Track 2 CPC+ practices. These are followed by detailed findings from the triple differences sensitivity test which is our key sensitivity test to assess potential bias due to differential regional effects of COVID-19.

The methods underlying our main impact analyses rely on a difference-in-differences estimation strategy that was adjusted to account for potential bias in our impact estimates due to the COVID-19 pandemic. In particular, we added COVID-19-specific region-level control variables to our regression models. Details on the additional control variables added to our models, and their specifications are described in Appendices 5.D (Implications of COVID-19 for the CPC+ Impact Evaluation) and 5.E (Empirical Strategy).

[1] In this appendix, we do not analyze or report on the practices that joined CPC+ in 2018, as these practices account for only 5 percent of the total number of practices participating in CPC+, and previous analyses found that the experiences of these practices were very similar to the experiences of those that joined CPC+ in 2017 (Anglin et al. 2020).

5.A.1. Medicare FFS service use

Table 5.A.1.1a. Regression-adjusted means and estimated impacts of CPC+ on selected Medicare service use outcomes for attributed Medicare FFS beneficiaries by program year and average across the five program years, Track 1

	Track 1 — Overall					
	CPC+ mean[a]	C mean[a]	Impact estimate[b] (SE)	Percentage impact[c]	90% confidence interval	*p*-Value
Service use (per 1,000 beneficiaries per year)						
Acute hospitalizations (short-stay acute care and critical access hospitals)						
Baseline	290	289	NA	NA	NA	NA
PY 1	289	288	-0.6 (1.5)	-0.2%	(-3.1, 1.9)	0.68
PY 2	285	285	-1.8 (1.6)	-0.6%	(-4.5, 0.9)	0.27
PY 3	284	286	-2.6 (1.8)	-0.9%	(-5.5, 0.3)	0.14
PY 4	243	247	-4.9*** (1.8)	-2.0%	(-7.8, -2.0)	0.01
PY 5	244	246	-2.6 (1.8)	-1.1%	(-5.6, 0.4)	0.15
PY 1 through 5	268	269	-2.5* (1.4)	-0.9%	(-4.9, -0.1)	0.08
Total ED visits, including observation stays[d]						
Baseline	711	710	NA	NA	NA	NA
PY 1	708	713	-6.7** (2.8)	-0.9%	(-11.4, -2.0)	0.02
PY 2	700	709	-10.9*** (3.2)	-1.5%	(-16.2, -5.7)	0.00
PY 3	699	712	-14.2*** (3.5)	-2.0%	(-19.9, -8.5)	0.00
PY 4	567	584	-18.7*** (3.9)	-3.2%	(-25.1, -12.4)	0.00
PY 5	604	624	-21.3*** (4.2)	-3.4%	(-28.2, -14.4)	0.00
PY 1 through 5	653	666	-14.3*** (3.0)	-2.1%	(-19.2, -9.5)	0.00
Outpatient ED visits, including observation stays						
Baseline	493	498	NA	NA	NA	NA
PY 1	490	501	-5.5** (2.3)	-1.1%	(-9.3, -1.7)	0.02
PY 2	484	497	-7.3*** (2.6)	-1.5%	(-11.6, -3.0)	0.01
PY 3	484	497	-8.1*** (2.9)	-1.6%	(-12.8, -3.3)	0.00
PY 4	376	392	-10.8*** (3.3)	-2.8%	(-16.3, -5.3)	0.00
PY 5	407	428	-15.7*** (3.7)	-3.7%	(-21.8, -9.7)	0.00
PY 1 through 5	446	461	-9.4*** (2.5)	-2.1%	(-13.5, -5.3)	0.00

Table 5.A.1.1a. ***(continued)***

	Track 1 — Overall					
	CPC+ mean[a]	C mean[a]	Impact estimate[b] (SE)	Percentage impact[c]	90% confidence interval	*p*-Value
Primary care substitutable outpatient ED visits[e]						
Baseline	192	195	NA	NA	NA	NA
PY 1	189	195	-2.2* (1.1)	-1.1%	(-4.1, -0.3)	0.06
PY 2	184	191	-3.8*** (1.3)	-2.0%	(-5.9, -1.6)	0.00
PY 3	181	190	-4.4*** (1.4)	-2.4%	(-6.7, -2.1)	0.00
PY 4	134	142	-4.6*** (1.6)	-3.3%	(-7.3, -1.9)	0.01
PY 5	135	146	-7.5*** (1.9)	-5.3%	(-10.6, -4.5)	0.00
PY 1 through 5	163	171	-4.5*** (1.2)	-2.7%	(-6.5, -2.5)	0.00
Potentially primary care preventable outpatient ED visits[e]						
Baseline	131	133	NA	NA	NA	NA
PY 1	129	133	-2.1** (0.8)	-1.6%	(-3.4, -0.7)	0.01
PY 2	127	130	-1.5 (0.9)	-1.1%	(-3.0, 0.1)	0.12
PY 3	126	130	-1.7* (0.9)	-1.3%	(-3.2, -0.1)	0.08
PY 4	97	101	-2.2** (1.1)	-2.3%	(-4.0, -0.5)	0.04
PY 5	101	106	-3.0*** (1.2)	-2.9%	(-4.9, -1.1)	0.01
PY 1 through 5	115	119	-2.1** (0.8)	-1.8%	(-3.4, -0.7)	0.01
Total Urgent Care Center (UCC) visits						
Baseline	104	111	NA	NA	NA	NA
PY 1	119	125	0.7 (1.7)	0.6%	(-2.1, 3.5)	0.67
PY 2	135	139	3.0 (2.6)	2.3%	(-1.3, 7.2)	0.25
PY 3	149	153	3.6 (3.8)	2.4%	(-2.6, 9.8)	0.34
PY 4	156	143	20.8*** (4.8)	15.4%	(12.9, 28.8)	0.00
PY 5	212	215	4.3 (6.2)	2.0%	(-5.9, 14.4)	0.49
PY 1 through 5	156	157	6.3** (3.2)	4.2%	(1.1, 11.6)	0.05
Primary care substitutable UCC visits						
Baseline	62	66	NA	NA	NA	NA
PY 1	72	75	0.1 (1.0)	0.2%	(-1.6, 1.8)	0.90
PY 2	82	83	1.7 (1.6)	2.1%	(-0.9, 4.3)	0.28
PY 3	90	91	3.0 (2.3)	3.4%	(-0.7, 6.7)	0.19
PY 4	104	92	15.8*** (3.0)	17.9%	(10.9, 20.8)	0.00
PY 5	83	82	4.2 (2.8)	5.3%	(-0.3, 8.7)	0.13
PY 1 through 5	87	85	4.8*** (1.8)	5.9%	(1.9, 7.8)	0.01

Table 5.A.1.1a. ***(continued)***

	Track 1 — Overall					
	CPC+ mean[a]	C mean[a]	Impact estimate[b] (SE)	Percentage impact[c]	90% confidence interval	*p*-Value
UCC visits that excludes COVID-related diagnoses						
Baseline	103	110	NA	NA	NA	NA
PY 1	118	125	0.4 (1.7)	0.4%	(-2.3, 3.2)	0.80
PY 2	134	138	2.9 (2.6)	2.2%	(-1.4, 7.1)	0.26
PY 3	148	152	3.6 (3.7)	2.5%	(-2.5, 9.7)	0.33
PY 4	106	106	7.2 (4.4)	7.3%	(-0.1, 14.5)	0.10
PY 5	121	130	-1.4 (4.2)	-1.1%	(-8.3, 5.5)	0.74
PY 1 through 5	125	130	2.6 (2.8)	2.1%	(-2.1, 7.2)	0.37
Ambulatory primary care visits (including to FQHCs, RHCs, and CAHs)[f]						
Baseline	4,255	4,370	NA	NA	NA	NA
PY 1	4,295	4,465	-54.3*** (15.1)	-1.2%	(-79.2, -29.4)	0.00
PY 2	4,340	4,474	-17.9 (19.2)	-0.4%	(-49.5, 13.6)	0.35
PY 3	4,406	4,521	0.3 (22.0)	0.0%	(-35.8, 36.4)	0.99
PY 4	3,991	4,126	-19.5 (27.4)	-0.5%	(-64.5, 25.6)	0.48
PY 5	4,244	4,360	0.1 (31.4)	0.0%	(-51.6, 51.8)	1.00
PY 1 through 5	4,252	4,385	-17.7 (19.9)	-0.4%	(-50.4, 15.1)	0.37
Ambulatory specialty care visits (including to FQHCs, RHCs, and CAHs)[f]						
Baseline	4,526	4,406	NA	NA	NA	NA
PY 1	4,474	4,347	7.8 (9.9)	0.2%	(-8.4, 24.0)	0.43
PY 2	4,496	4,353	23.7* (12.8)	0.5%	(2.7, 44.8)	0.06
PY 3	4,403	4,269	13.9 (14.8)	0.3%	(-10.4, 38.2)	0.35
PY 4	3,830	3,708	1.5 (17.0)	0.0%	(-26.5, 29.4)	0.93
PY 5	4,182	4,070	-8.4 (18.9)	-0.2%	(-39.4, 22.6)	0.66
PY 1 through 5	4,266	4,138	8.1 (12.8)	0.2%	(-12.9, 29.2)	0.53
Unweighted sample sizes for measures per 1,000 beneficiaries per year[g]						
Number of practices	1,373	5,243				
Number of beneficiaries	1,549,585	5,347,499				
Number of beneficiary-years	5,916,394	20,150,090				

Source: Mathematica's analysis of Medicare claims data from January 2013 through December 2021.

Notes: This table indicates which estimates are statistically significant; when we interpret evidence, we combine evidence from the magnitude of the effect, the *p*-values, findings on related outcomes, subgroups, sensitivity tests, and other data sources about model implementation.

[a] We report the actual, unadjusted averages in the baseline period which are similar for the CPC+ and comparison groups due to matching. In the intervention periods, the comparison group mean is computed by subtracting the regression adjusted difference between the CPC+ and comparison means in each time period from the CPC+ mean in that same time period.

Table 5.A.1.1a. ***(continued)***

[b] Each impact estimate is regression-adjusted using a difference-in-differences analysis that reflects the difference of the average outcome for Medicare FFS beneficiaries attributed to CPC+ practices in the five years of CPC+ to the average outcome in the baseline year, relative to the same difference over time for Medicare FFS beneficiaries attributed to comparison practices, while controlling for beneficiary characteristics and practice fixed effects.

[c] We calculated percentage impacts relative to what the CPC+ mean would have been in Program Years 1 through 5 (separately and combined) in the absence of the intervention—that is, the unadjusted CPC+ mean minus the impact estimate.

[d] Total ED visits include ED/observation stays that led to a hospitalization, including a psychiatric hospitalization.

[e] The sum of primary care substitutable outpatient ED visits and potentially primary care preventable outpatient ED visits is less than total outpatient ED visits because total outpatient ED visits include those for other care needs, such as injuries, mental health, drugs, and alcohol.

[f] Ambulatory visits with primary care practitioners and specialists include office-based visits and visits at home, as well as visits in other settings, such as FQHCs, RHCs, and CAHs.

[g] After accounting for weights that adjust for matching and time observed in Medicare FFS, the effective sample sizes fall but are still substantial. For the comparison group, the effective sample size is 45 percent of the size of the actual comparison group. The effective sample size for the CPC+ group is 96 percent of the actual sample size because it is affected only by time observed (and not by the matching weights).

*/**/*** Significantly different from zero at the 0.10/0.05/0.01 level, two-tailed test.

C = comparison; CAH = critical access hospital; ED = emergency department; FFS = fee-for-service; FQHC = federally qualified health center; NA = not applicable; PY = Program Year; RHC = rural health clinic; SE = standard error.

Table 5.A.1.1b. Regression-adjusted means and estimated impacts of CPC+ on selected Medicare service use outcomes for attributed Medicare FFS beneficiaries by program year and average across the five program years, Track 1 by SSP status

	Track 1 — SSP						Track 1 — Non-SSP						
	CPC+ mean[a]	C mean[a]	Impact estimate[b] (SE)	Percentage impact[c]	90% confidence interval	*p*-Value	CPC+ mean[a]	C mean[a]	Impact estimate[b] (SE)	Percentage impact[c]	90% confidence interval	*p*-Value	*p*-Value for SSP vs. non-SSP difference
Service use (per 1,000 beneficiaries per year)													
Acute hospitalizations (short-stay acute care and critical access hospitals)													
Baseline	291	289	NA	NA	NA	NA	289	288	NA	NA	NA	NA	NA
PY 1	289	290	-2.7 (1.9)	-0.9%	(-5.8, 0.5)	0.16	289	286	1.6 (2.4)	0.6%	(-2.3, 5.5)	0.51	0.16
PY 2	286	287	-2.2 (2.1)	-0.8%	(-5.7, 1.3)	0.30	283	283	-1.4 (2.6)	-0.5%	(-5.6, 2.8)	0.57	0.82
PY 3	286	289	-4.9** (2.2)	-1.7%	(-8.6, -1.2)	0.03	283	282	0.0 (2.8)	0.0%	(-4.6, 4.5)	1.00	0.17
PY 4	245	251	-8.0*** (2.3)	-3.2%	(-11.9, -4.2)	0.00	241	242	-1.4 (2.8)	-0.6%	(-5.9, 3.2)	0.62	0.07
PY 5	250	253	-5.1** (2.5)	-2.0%	(-9.2, -1.0)	0.04	239	238	0.7 (2.7)	0.3%	(-3.7, 5.0)	0.81	0.12
PY 1 through 5	270	273	-4.5** (1.8)	-1.6%	(-7.5, -1.5)	0.01	266	265	-0.1 (2.2)	0.0%	(-3.8, 3.6)	0.96	0.13
Total ED visits, including observation stays[d]													
Baseline	698	696	NA	NA	NA	NA	724	724	NA	NA	NA	NA	NA
PY 1	696	701	-7.6** (3.8)	-1.1%	(-13.8, -1.4)	0.04	721	726	-5.6 (4.3)	-0.8%	(-12.6, 1.5)	0.20	0.72
PY 2	688	696	-11.0*** (4.2)	-1.6%	(-18.0, -4.1)	0.01	713	724	-10.8** (4.7)	-1.5%	(-18.6, -3.1)	0.02	0.98
PY 3	689	701	-14.9*** (4.5)	-2.1%	(-22.2, -7.6)	0.00	710	724	-13.3** (5.4)	-1.8%	(-22.1, -4.5)	0.01	0.82
PY 4	558	577	-21.8*** (5.1)	-3.8%	(-30.2, -13.3)	0.00	577	591	-14.1** (6.0)	-2.4%	(-23.9, -4.2)	0.02	0.33
PY 5	601	621	-23.0*** (5.7)	-3.7%	(-32.4, -13.7)	0.00	608	627	-19.2*** (6.1)	-3.1%	(-29.2, -9.1)	0.00	0.64
PY 1 through 5	644	657	-15.4*** (3.9)	-2.3%	(-21.9, -8.9)	0.00	663	675	-12.7*** (4.4)	-1.9%	(-19.9, -5.4)	0.00	0.64

Table 5.A.1.1b. ***(continued)***

	Track 1 — SSP						Track 1 — Non-SSP						
	CPC+ mean[a]	C mean[a]	Impact estimate[b] (SE)	Percentage impact[c]	90% confidence interval	*p*-Value	CPC+ mean[a]	C mean[a]	Impact estimate[b] (SE)	Percentage impact[c]	90% confidence interval	*p*-Value	*p*-Value for SSP vs. non-SSP difference
Outpatient ED visits, including observation stays													
Baseline	476	480	NA	NA	NA	NA	510	518	NA	NA	NA	NA	NA
PY 1	475	484	-5.6* (3.0)	-1.2%	(-10.5, -0.6)	0.06	506	520	-5.2 (3.5)	-1.0%	(-11.0, 0.5)	0.13	0.94
PY 2	467	479	-8.2** (3.5)	-1.7%	(-13.9, -2.5)	0.02	502	516	-6.3 (4.0)	-1.2%	(-12.9, 0.2)	0.11	0.72
PY 3	469	480	-7.5** (3.6)	-1.6%	(-13.5, -1.5)	0.04	499	516	-8.7* (4.5)	-1.7%	(-16.0, -1.3)	0.05	0.84
PY 4	361	377	-13.6*** (4.4)	-3.6%	(-20.9, -6.3)	0.00	392	407	-6.6 (5.1)	-1.7%	(-15.0, 1.9)	0.20	0.30
PY 5	395	416	-17.6*** (4.8)	-4.3%	(-25.4, -9.8)	0.00	419	441	-14.0** (5.5)	-3.2%	(-23.0, -4.9)	0.01	0.62
PY 1 through 5	432	445	-10.2*** (3.3)	-2.3%	(-15.6, -4.8)	0.00	461	478	-8.2** (3.8)	-1.8%	(-14.4, -2.0)	0.03	0.69
Primary care substitutable outpatient ED visits[e]													
Baseline	186	187	NA	NA	NA	NA	198	204	NA	NA	NA	NA	NA
PY 1	183	187	-3.0** (1.5)	-1.6%	(-5.5, -0.6)	0.04	196	203	-1.2 (1.8)	-0.6%	(-4.1, 1.7)	0.50	0.42
PY 2	178	183	-4.4*** (1.7)	-2.4%	(-7.2, -1.7)	0.01	191	200	-3.0 (2.0)	-1.5%	(-6.2, 0.2)	0.13	0.57
PY 3	176	182	-4.2** (1.8)	-2.3%	(-7.1, -1.3)	0.02	187	198	-4.6** (2.2)	-2.4%	(-8.3, -1.0)	0.04	0.88
PY 4	128	135	-5.6*** (2.1)	-4.2%	(-9.2, -2.1)	0.01	139	148	-3.0 (2.6)	-2.1%	(-7.2, 1.2)	0.24	0.42
PY 5	131	140	-8.4*** (2.3)	-6.1%	(-12.2, -4.7)	0.00	139	151	-6.1** (2.8)	-4.2%	(-10.7, -1.5)	0.03	0.52
PY 1 through 5	158	165	-5.0*** (1.6)	-3.1%	(-7.7, -2.4)	0.00	169	179	-3.6* (1.9)	-2.1%	(-6.8, -0.5)	0.06	0.57
Potentially primary care preventable outpatient ED visits[e]													
Baseline	125	127	NA	NA	NA	NA	138	140	NA	NA	NA	NA	NA
PY 1	123	127	-1.4 (1.1)	-1.1%	(-3.1, 0.3)	0.19	134	139	-2.7** (1.3)	-2.0%	(-4.8, -0.6)	0.03	0.43
PY 2	121	124	-0.8 (1.2)	-0.7%	(-2.8, 1.2)	0.50	133	137	-2.1 (1.4)	-1.5%	(-4.4, 0.3)	0.14	0.50
PY 3	121	124	-0.7 (1.2)	-0.6%	(-2.7, 1.2)	0.53	132	136	-2.6* (1.5)	-1.9%	(-5.0, -0.2)	0.08	0.33
PY 4	92	96	-2.2 (1.4)	-2.4%	(-4.5, 0.0)	0.11	102	106	-2.1 (1.7)	-2.0%	(-4.8, 0.6)	0.21	0.94
PY 5	97	102	-2.6* (1.5)	-2.6%	(-5.0, -0.2)	0.07	105	111	-3.4* (1.8)	-3.1%	(-6.3, -0.5)	0.06	0.73
PY 1 through 5	110	114	-1.5 (1.1)	-1.3%	(-3.2, 0.2)	0.15	120	125	-2.6** (1.3)	-2.1%	(-4.7, -0.5)	0.04	0.52

Table 5.A.1.1b. ***(continued)***

	Track 1 — SSP						Track 1 — Non-SSP						
	CPC+ mean[a]	C mean[a]	Impact estimate[b] (SE)	Percentage impact[c]	90% confidence interval	*p*-Value	CPC+ mean[a]	C mean[a]	Impact estimate[b] (SE)	Percentage impact[c]	90% confidence interval	*p*-Value	*p*-Value for SSP vs. non-SSP difference
Total Urgent Care Center (UCC) visits													
Baseline	114	112	NA	NA	NA	NA	93	109	NA	NA	NA	NA	NA
PY 1	132	128	1.9 (2.3)	1.5%	(-1.8, 5.6)	0.40	105	122	-0.6 (2.5)	-0.5%	(-4.7, 3.6)	0.83	0.47
PY 2	151	142	6.7** (2.9)	4.7%	(2.0, 11.5)	0.02	118	135	-1.1 (4.4)	-0.9%	(-8.3, 6.2)	0.80	0.14
PY 3	167	161	3.8 (4.4)	2.3%	(-3.5, 11.0)	0.39	131	144	3.6 (6.2)	2.8%	(-6.7, 13.9)	0.56	0.98
PY 4	179	156	21.7*** (5.4)	13.8%	(12.8, 30.6)	0.00	133	128	21.4** (8.3)	19.2%	(7.7, 35.1)	0.01	0.98
PY 5	251	252	-2.8 (9.1)	-1.1%	(-17.7, 12.1)	0.76	173	180	9.4 (8.4)	5.7%	(-4.4, 23.2)	0.26	0.33
PY 1 through 5	178	169	6.3* (3.8)	3.7%	(0.1, 12.5)	0.10	133	144	6.3 (5.1)	5.0%	(-2.2, 14.8)	0.22	1.00
Primary care substitutable UCC visits													
Baseline	68	67	NA	NA	NA	NA	56	64	NA	NA	NA	NA	NA
PY 1	79	77	1.1 (1.4)	1.4%	(-1.3, 3.4)	0.45	64	73	-0.9 (1.5)	-1.3%	(-3.4, 1.7)	0.58	0.35
PY 2	91	86	4.3** (1.9)	5.0%	(1.3, 7.4)	0.02	72	81	-1.0 (2.6)	-1.4%	(-5.4, 3.3)	0.70	0.10
PY 3	101	96	4.3 (2.7)	4.5%	(-0.2, 8.8)	0.11	79	85	1.7 (3.6)	2.2%	(-4.2, 7.7)	0.63	0.57
PY 4	121	102	18.6*** (3.6)	18.1%	(12.6, 24.6)	0.00	87	81	13.8*** (5.0)	19.0%	(5.7, 22.0)	0.01	0.45
PY 5	94	88	5.3 (3.6)	6.0%	(-0.7, 11.3)	0.15	72	76	3.4 (4.2)	5.0%	(-3.5, 10.3)	0.41	0.73
PY 1 through 5	98	90	6.6*** (2.2)	7.2%	(3.0, 10.1)	0.00	75	80	3.2 (2.9)	4.5%	(-1.5, 8.0)	0.27	0.36
UCC visits that excludes COVID-related diagnoses													
Baseline	113	112	NA	NA	NA	NA	92	108	NA	NA	NA	NA	NA
PY 1	131	128	1.5 (2.2)	1.2%	(-2.2, 5.2)	0.49	104	121	-0.7 (2.5)	-0.7%	(-4.8, 3.5)	0.78	0.51
PY 2	150	141	6.7** (2.9)	4.7%	(2.0, 11.4)	0.02	117	134	-1.1 (4.4)	-0.9%	(-8.4, 6.1)	0.80	0.14
PY 3	166	160	4.1 (4.4)	2.6%	(-3.1, 11.3)	0.34	130	143	3.4 (6.1)	2.7%	(-6.6, 13.5)	0.58	0.92
PY 4	117	112	2.6 (4.5)	2.3%	(-4.8, 10.0)	0.56	95	98	13.1 (8.0)	16.0%	(0.0, 26.3)	0.10	0.25
PY 5	132	133	-2.6 (5.5)	-1.9%	(-11.7, 6.4)	0.64	110	124	2.1 (6.3)	1.9%	(-8.3, 12.5)	0.74	0.58
PY 1 through 5	139	135	2.7 (3.2)	2.0%	(-2.6, 8.0)	0.40	111	124	3.2 (4.7)	3.0%	(-4.5, 11.0)	0.49	0.93

Table 5.A.1.1b. ***(continued)***

	Track 1 — SSP						Track 1 — Non-SSP						
	CPC+ mean[a]	C mean[a]	Impact estimate[b] (SE)	Percentage impact[c]	90% confidence interval	*p*-Value	CPC+ mean[a]	C mean[a]	Impact estimate[b] (SE)	Percentage impact[c]	90% confidence interval	*p*-Value	*p*-Value for SSP vs. non-SSP difference
Ambulatory primary care visits (including to FQHCs, RHCs, and CAHs)[f]													
Baseline	4,207	4,341	NA	NA	NA	NA	4,305	4,403	NA	NA	NA	NA	NA
PY 1	4,260	4,438	-44.9** (18.3)	-1.0%	(-75.0, -14.8)	0.01	4,332	4,494	-64.0*** (24.5)	-1.5%	(-104.3, -23.7)	0.01	0.53
PY 2	4,297	4,432	-2.3 (24.6)	-0.1%	(-42.7, 38.1)	0.93	4,386	4,519	-34.8 (29.8)	-0.8%	(-83.9, 14.3)	0.24	0.40
PY 3	4,362	4,490	5.8 (28.3)	0.1%	(-40.8, 52.3)	0.84	4,451	4,554	-5.1 (33.9)	-0.1%	(-61.0, 50.7)	0.88	0.81
PY 4	3,956	4,095	-5.7 (32.8)	-0.1%	(-59.6, 48.3)	0.86	4,026	4,155	-30.6 (44.4)	-0.8%	(-103.6, 42.4)	0.49	0.65
PY 5	4,218	4,354	-2.6 (39.4)	-0.1%	(-67.5, 62.2)	0.95	4,271	4,376	-7.0 (49.8)	-0.2%	(-88.9, 75.0)	0.89	0.95
PY 1 through 5	4,216	4,359	-9.5 (24.5)	-0.2%	(-49.9, 30.8)	0.70	4,289	4,415	-27.5 (31.6)	-0.6%	(-79.5, 24.6)	0.39	0.66
Ambulatory specialty care visits (including to FQHCs, RHCs, and CAHs)[f]													
Baseline	4,836	4,610	NA	NA	NA	NA	4,200	4,183	NA	NA	NA	NA	NA
PY 1	4,765	4,549	-9.8 (13.0)	-0.2%	(-31.2, 11.6)	0.45	4,166	4,123	26.3* (14.9)	0.6%	(1.9, 50.8)	0.08	0.07
PY 2	4,818	4,571	21.0 (17.1)	0.4%	(-7.1, 49.1)	0.22	4,157	4,111	28.4 (18.9)	0.7%	(-2.7, 59.4)	0.13	0.77
PY 3	4,735	4,501	7.6 (19.9)	0.2%	(-25.1, 40.4)	0.70	4,057	4,016	23.4 (21.3)	0.6%	(-11.6, 58.5)	0.27	0.59
PY 4	4,118	3,927	-35.4 (22.8)	-0.9%	(-72.9, 2.1)	0.12	3,535	3,464	53.0** (24.4)	1.5%	(12.8, 93.3)	0.03	0.01
PY 5	4,533	4,345	-38.1 (25.8)	-0.8%	(-80.6, 4.4)	0.14	3,823	3,777	28.8 (27.6)	0.8%	(-16.6, 74.2)	0.30	0.08
PY 1 through 5	4,585	4,368	-9.2 (17.1)	-0.2%	(-37.2, 18.9)	0.59	3,935	3,885	31.6* (18.7)	0.8%	(0.8, 62.3)	0.09	0.11
Unweighted sample sizes for measures per 1,000 beneficiaries per year[g]													
Number of practices	738	2,979					635	2,264					
Number of beneficiaries	798,817	3,129,830					753,337	2,233,041					
Number of beneficiary-years	3,017,546	11,762,356					2,898,848	8,387,734					

Source: Mathematica's analysis of Medicare claims data from January 2013 through December 2021.

Notes: This table indicates which estimates are statistically significant; when we interpret evidence, we combine evidence from the magnitude of the effect, the *p*-values, findings on related outcomes, subgroups, sensitivity tests, and other data sources about model implementation.

[a] We report the actual, unadjusted averages in the baseline period which are similar for the CPC+ and comparison groups due to matching. In the intervention periods, the comparison group mean is computed by subtracting the regression adjusted difference between the CPC+ and comparison means in each time period from the CPC+ mean in that same time period.

[b] Each impact estimate is regression-adjusted using a difference-in-differences analysis that reflects the difference of the average outcome for Medicare FFS beneficiaries attributed to CPC+ practices in the five years of CPC+ to the average outcome in the baseline year, relative to the same difference over time for Medicare FFS beneficiaries attributed to comparison practices, while controlling for beneficiary characteristics and practice fixed effects.

Table 5.A.1.1b. ***(continued)***

[c] We calculated percentage impacts relative to what the CPC+ mean would have been in Program Years 1 through 5 (separately and combined) in the absence of the intervention—that is, the unadjusted CPC+ mean minus the impact estimate.

[d] Total ED visits include ED/observation stays that led to a hospitalization, including a psychiatric hospitalization.

[e] The sum of primary care substitutable outpatient ED visits and potentially primary care preventable outpatient ED visits is less than total outpatient ED visits because total outpatient ED visits include those for other care needs, such as injuries, mental health, drugs, and alcohol.

[f] Ambulatory visits with primary care practitioners and specialists include office-based visits and visits at home, as well as visits in other settings, such as FQHCs, RHCs, and CAHs.

[g] After accounting for weights that adjust for matching and time observed in Medicare FFS, the effective sample sizes fall but are still substantial. For the comparison group, the effective sample size is 43 to 50 percent of the size of the actual comparison group. The effective sample size for the CPC+ group is 96 percent of the actual sample size because it is affected only by time observed (and not by the matching weights).

*/**/*** Significantly different from zero at the 0.10/0.05/0.01 level, two-tailed test.

C = comparison; CAH = critical access hospital; ED = emergency department; FFS = fee-for-service; FQHC = federally qualified health center; NA = not applicable; PY = Program Year; RHC = rural health clinic; SE = standard error; SSP = Medicare Shared Savings Program.

Table 5.A.1.2a. Regression-adjusted means and estimated impacts of CPC+ on selected Medicare service use outcomes for attributed Medicare FFS beneficiaries by program year and average across the five program years, Track 2

	Track 2 — Overall					
	CPC+ mean[a]	C mean[a]	Impact estimate[b] (SE)	Percentage impact[c]	90% confidence interval	*p*-Value
Service use (per 1,000 beneficiaries per year)						
Acute hospitalizations (short-stay acute care and critical access hospitals)						
Baseline	292	288	NA	NA	NA	NA
PY 1	292	289	-0.5 (1.6)	-0.2%	(-3.2, 2.1)	0.74
PY 2	289	286	-1.5 (1.7)	-0.5%	(-4.3, 1.3)	0.38
PY 3	286	287	-4.8*** (1.8)	-1.7%	(-7.8, -1.8)	0.01
PY 4	245	246	-4.8** (1.9)	-1.9%	(-8.0, -1.7)	0.01
PY 5	246	243	-1.9 (2.0)	-0.8%	(-5.1, 1.3)	0.34
PY 1 through 5	270	269	-2.7* (1.6)	-1.0%	(-5.3, -0.2)	0.08
Total ED visits, including observation stays[d]						
Baseline	710	705	NA	NA	NA	NA
PY 1	705	708	-8.1*** (3.0)	-1.1%	(-13.0, -3.2)	0.01
PY 2	702	706	-10.0*** (3.3)	-1.4%	(-15.5, -4.6)	0.00
PY 3	699	708	-14.5*** (3.7)	-2.0%	(-20.6, -8.3)	0.00
PY 4	570	579	-14.8*** (4.2)	-2.5%	(-21.7, -7.9)	0.00
PY 5	605	614	-14.3*** (4.3)	-2.3%	(-21.4, -7.2)	0.00
PY 1 through 5	653	660	-12.4*** (3.2)	-1.9%	(-17.7, -7.1)	0.00
Outpatient ED visits, including observation stays						
Baseline	492	492	NA	NA	NA	NA
PY 1	486	494	-7.8*** (2.4)	-1.6%	(-11.6, -3.9)	0.00
PY 2	483	490	-6.6** (2.7)	-1.4%	(-11.0, -2.3)	0.01
PY 3	483	491	-7.5** (3.0)	-1.5%	(-12.4, -2.5)	0.01
PY 4	378	386	-8.3** (3.5)	-2.2%	(-14.1, -2.5)	0.02
PY 5	408	419	-11.2*** (3.6)	-2.7%	(-17.2, -5.2)	0.00
PY 1 through 5	445	454	-8.3*** (2.6)	-1.8%	(-12.6, -3.9)	0.00
Primary care substitutable outpatient ED visits[e]						
Baseline	191	192	NA	NA	NA	NA
PY 1	187	192	-4.2*** (1.2)	-2.2%	(-6.1, -2.3)	0.00
PY 2	183	189	-4.4*** (1.3)	-2.3%	(-6.4, -2.3)	0.00
PY 3	181	187	-5.7*** (1.4)	-3.0%	(-8.0, -3.3)	0.00
PY 4	134	140	-5.3*** (1.7)	-3.8%	(-8.1, -2.6)	0.00
PY 5	135	142	-6.8*** (1.7)	-4.8%	(-9.6, -4.0)	0.00

Table 5.A.1.2a. ***(continued)***

	Track 2 — Overall					
	CPC+ mean[a]	C mean[a]	Impact estimate[b] (SE)	Percentage impact[c]	90% confidence interval	*p*-Value
PY 1 through 5	163	169	-5.3*** (1.2)	-3.1%	(-7.3, -3.2)	0.00
Potentially primary care preventable outpatient ED visits[e]						
Baseline	133	131	NA	NA	NA	NA
PY 1	130	131	-2.5*** (0.8)	-1.9%	(-3.8, -1.2)	0.00
PY 2	127	128	-2.3** (0.9)	-1.8%	(-3.8, -0.8)	0.01
PY 3	127	128	-2.7*** (1.0)	-2.1%	(-4.3, -1.1)	0.01
PY 4	97	99	-2.9*** (1.1)	-2.9%	(-4.6, -1.1)	0.01
PY 5	102	104	-3.4*** (1.1)	-3.3%	(-5.2, -1.7)	0.00
PY 1 through 5	116	117	-2.8*** (0.8)	-2.3%	(-4.1, -1.4)	0.00
Total Urgent Care Center (UCC) visits						
Baseline	97	106	NA	NA	NA	NA
PY 1	111	119	1.1 (2.1)	1.0%	(-2.3, 4.6)	0.58
PY 2	124	130	2.2 (2.9)	1.8%	(-2.6, 7.1)	0.44
PY 3	134	145	-2.3 (3.6)	-1.7%	(-8.2, 3.7)	0.53
PY 4	136	135	9.0** (3.9)	7.1%	(2.5, 15.4)	0.02
PY 5	186	189	5.6 (5.8)	3.1%	(-3.9, 15.1)	0.33
PY 1 through 5	140	145	3.0 (3.1)	2.2%	(-2.1, 8.0)	0.33
Primary care substitutable UCC visits						
Baseline	58	62	NA	NA	NA	NA
PY 1	67	71	0.5 (1.3)	0.7%	(-1.6, 2.6)	0.70
PY 2	74	78	0.8 (1.7)	1.1%	(-2.0, 3.6)	0.64
PY 3	81	85	-0.7 (2.1)	-0.9%	(-4.2, 2.8)	0.75
PY 4	88	85	7.1*** (2.4)	8.7%	(3.1, 11.1)	0.00
PY 5	76	75	5.5** (2.5)	7.8%	(1.4, 9.7)	0.03
PY 1 through 5	78	79	2.5 (1.7)	3.4%	(-0.3, 5.4)	0.14
UCC visits that excludes COVID-related diagnoses						
Baseline	97	105	NA	NA	NA	NA
PY 1	111	118	1.0 (2.1)	0.9%	(-2.4, 4.4)	0.64
PY 2	124	130	2.2 (2.9)	1.8%	(-2.5, 7.0)	0.44
PY 3	133	144	-2.1 (3.6)	-1.6%	(-8.1, 3.8)	0.55
PY 4	97	103	2.1 (3.6)	2.2%	(-3.8, 8.0)	0.56
PY 5	114	121	1.5 (4.0)	1.4%	(-5.1, 8.2)	0.70
PY 1 through 5	116	123	0.9 (2.8)	0.8%	(-3.7, 5.5)	0.75

Table 5.A.1.2a. ***(continued)***

	Track 2 — Overall					
	CPC+ mean[a]	C mean[a]	Impact estimate[b] (SE)	Percentage impact[c]	90% confidence interval	*p*-Value
Ambulatory primary care visits (including to FQHCs, RHCs, and CAHs)[f]						
Baseline	4,361	4,438	NA	NA	NA	NA
PY 1	4,364	4,513	-71.3*** (16.3)	-1.6%	(-98.1, -44.6)	0.00
PY 2	4,393	4,516	-45.8** (21.5)	-1.0%	(-81.1, -10.4)	0.03
PY 3	4,449	4,561	-35.5 (26.5)	-0.8%	(-79.0, 8.1)	0.18
PY 4	4,019	4,124	-28.3 (27.8)	-0.7%	(-74.0, 17.5)	0.31
PY 5	4,236	4,355	-42.7 (32.7)	-1.0%	(-96.5, 11.2)	0.19
PY 1 through 5	4,286	4,408	-44.4** (21.9)	-1.0%	(-80.5, -8.3)	0.04
Ambulatory specialty care visits (including to FQHCs, RHCs, and CAHs)[f]						
Baseline	4,425	4,322	NA	NA	NA	NA
PY 1	4,380	4,279	-2.5 (10.3)	-0.1%	(-19.5, 14.4)	0.81
PY 2	4,362	4,272	-12.5 (14.5)	-0.3%	(-36.4, 11.4)	0.39
PY 3	4,270	4,187	-19.5 (16.6)	-0.5%	(-46.9, 7.8)	0.24
PY 4	3,695	3,620	-28.7 (18.6)	-0.8%	(-59.4, 1.9)	0.12
PY 5	4,019	3,971	-55.3*** (20.3)	-1.4%	(-88.8, -21.9)	0.01
PY 1 through 5	4,132	4,052	-23.7* (14.1)	-0.6%	(-47.0, -0.4)	0.09
Unweighted sample sizes for measures per 1,000 beneficiaries per year[g]						
Number of practices	1,515	3,783				
Number of beneficiaries	1,896,880	4,507,499				
Number of beneficiary-years	7,225,289	17,054,519				

Source: Mathematica's analysis of Medicare claims data from January 2013 through December 2021.

Notes: This table indicates which estimates are statistically significant; when we interpret evidence, we combine evidence from the magnitude of the effect, the *p*-values, findings on related outcomes, subgroups, sensitivity tests, and other data sources about model implementation.

[a] We report the actual, unadjusted averages in the baseline period which are similar for the CPC+ and comparison groups due to matching. In the intervention periods, the comparison group mean is computed by subtracting the regression adjusted difference between the CPC+ and comparison means in each time period from the CPC+ mean in that same time period.

[b] Each impact estimate is regression-adjusted using a difference-in-differences analysis that reflects the difference of the average outcome for Medicare FFS beneficiaries attributed to CPC+ practices in the five years of CPC+ to the average outcome in the baseline year, relative to the same difference over time for Medicare FFS beneficiaries attributed to comparison practices, while controlling for beneficiary characteristics and practice fixed effects.

[c] We calculated percentage impacts relative to what the CPC+ mean would have been in Program Years 1 through 5 (separately and combined) in the absence of the intervention—that is, the unadjusted CPC+ mean minus the impact estimate.

[d] Total ED visits include ED/observation stays that led to a hospitalization, including a psychiatric hospitalization.

[e] The sum of primary care substitutable outpatient ED visits and potentially primary care preventable outpatient ED visits is less than total outpatient ED visits because total outpatient ED visits include those for other care needs, such as injuries, mental health, drugs, and alcohol.

[f] Ambulatory visits with primary care practitioners and specialists include office-based visits and visits at home, as well as visits in other settings, such as FQHCs, RHCs, and CAHs.

[g] After accounting for weights that adjust for matching and time observed in Medicare FFS, the effective sample sizes fall but are still substantial. For the comparison group, the effective sample size is 40 percent of the size of the actual comparison group. The effective sample size for the CPC+ group is 96 percent of the actual sample size because it is affected only by time observed (and not by the matching weights).

*/**/*** Significantly different from zero at the 0.10/0.05/0.01 level, two-tailed test.

C = comparison; CAH = critical access hospital; ED = emergency department; FFS = fee-for-service; FQHC = federally qualified health center; NA = not applicable; PY = Program Year; RHC = rural health clinic; SE = standard error.

Table 5.A.1.2b. Regression-adjusted means and estimated impacts of CPC+ on selected Medicare service use outcomes for attributed Medicare FFS beneficiaries by program year and average across the five program years, Track 2 by SSP status

	Track 2 — SSP						Track 2 — Non-SSP						
	CPC+ mean[a]	C mean[a]	Impact estimate[b] (SE)	Percentage impact[c]	90% confidence interval	*p*-Value	CPC+ mean[a]	C mean[a]	Impact estimate[b] (SE)	Percentage impact[c]	90% confidence interval	*p*-Value	*p*-Value for SSP vs. non-SSP difference
Service use (per 1,000 beneficiaries per year)													
Acute hospitalizations (short-stay acute care and critical access hospitals)													
Baseline	300	291	NA	NA	NA	NA	287	286	NA	NA	NA	NA	NA
PY 1	302	293	-0.4 (2.4)	-0.1%	(-4.3, 3.4)	0.85	285	285	-0.6 (2.2)	-0.2%	(-4.1, 3.0)	0.80	0.97
PY 2	297	289	-0.1 (2.6)	0.0%	(-4.2, 4.1)	0.98	282	284	-2.7 (2.3)	-0.9%	(-6.5, 1.1)	0.25	0.44
PY 3	296	290	-2.1 (2.7)	-0.7%	(-6.6, 2.5)	0.45	278	285	-7.0*** (2.4)	-2.5%	(-11.0, -3.0)	0.00	0.18
PY 4	253	248	-3.9 (3.0)	-1.5%	(-8.9, 1.1)	0.19	239	243	-5.1** (2.3)	-2.1%	(-8.9, -1.2)	0.03	0.76
PY 5	256	250	-2.3 (3.1)	-0.9%	(-7.5, 2.9)	0.46	237	238	-1.9 (2.4)	-0.8%	(-5.9, 2.1)	0.44	0.90
PY 1 through 5	280	273	-1.7 (2.4)	-0.6%	(-5.7, 2.2)	0.47	263	266	-3.5* (2.0)	-1.3%	(-6.8, -0.3)	0.08	0.57
Total ED visits, including observation stays[d]													
Baseline	705	692	NA	NA	NA	NA	715	715	NA	NA	NA	NA	NA
PY 1	700	696	-8.0* (4.3)	-1.1%	(-15.1, -1.0)	0.06	709	717	-8.0* (4.1)	-1.1%	(-14.8, -1.2)	0.05	1.00
PY 2	695	692	-8.7* (4.7)	-1.2%	(-16.5, -1.0)	0.06	707	718	-11.1** (4.6)	-1.5%	(-18.7, -3.4)	0.02	0.72
PY 3	694	695	-13.2** (5.6)	-1.9%	(-22.5, -3.9)	0.02	704	719	-15.5*** (5.0)	-2.2%	(-23.8, -7.3)	0.00	0.76
PY 4	562	574	-24.2*** (6.3)	-4.1%	(-34.5, -13.9)	0.00	576	580	-4.9 (5.4)	-0.8%	(-13.8, 4.1)	0.37	0.02
PY 5	601	609	-20.7*** (6.6)	-3.3%	(-31.6, -9.8)	0.00	608	614	-6.6 (5.5)	-1.1%	(-15.6, 2.4)	0.23	0.10
PY 1 through 5	648	650	-14.8*** (4.8)	-2.2%	(-22.7, -6.9)	0.00	658	667	-9.5** (4.3)	-1.4%	(-16.5, -2.5)	0.03	0.41
Outpatient ED visits, including observation stays													
Baseline	479	475	NA	NA	NA	NA	502	506	NA	NA	NA	NA	NA
PY 1	471	476	-9.3*** (3.4)	-1.9%	(-14.9, -3.7)	0.01	498	508	-6.5** (3.2)	-1.3%	(-11.8, -1.2)	0.04	0.55
PY 2	468	472	-8.3** (3.8)	-1.7%	(-14.5, -2.1)	0.03	496	505	-5.4 (3.7)	-1.1%	(-11.5, 0.8)	0.15	0.59
PY 3	469	472	-7.8* (4.3)	-1.6%	(-15.0, -0.7)	0.07	495	506	-7.2* (4.1)	-1.4%	(-14.0, -0.4)	0.08	0.92

Table 5.A.1.2b. ***(continued)***

	Track 2 — SSP						Track 2 — Non-SSP						
	CPC+ mean[a]	C mean[a]	Impact estimate[b] (SE)	Percentage impact[c]	90% confidence interval	*p*-Value	CPC+ mean[a]	C mean[a]	Impact estimate[b] (SE)	Percentage impact[c]	90% confidence interval	*p*-Value	*p*-Value for SSP vs. non-SSP difference
PY 4	362	377	-19.2*** (5.2)	-5.0%	(-27.9, -10.6)	0.00	391	392	2.3 (4.7)	0.6%	(-5.4, 9.9)	0.63	0.00
PY 5	392	408	-20.1*** (5.4)	-4.9%	(-29.0, -11.3)	0.00	420	425	-1.3 (4.6)	-0.3%	(-8.9, 6.4)	0.78	0.01
PY 1 through 5	430	438	-12.7*** (3.8)	-2.9%	(-19.0, -6.4)	0.00	457	465	-3.8 (3.5)	-0.8%	(-9.6, 1.9)	0.28	0.09
Primary care substitutable outpatient ED visits[e]													
Baseline	186	185	NA	NA	NA	NA	195	197	NA	NA	NA	NA	NA
PY 1	181	185	-5.7*** (1.7)	-3.1%	(-8.5, -2.9)	0.00	192	197	-3.0* (1.6)	-1.5%	(-5.6, -0.4)	0.06	0.25
PY 2	177	181	-6.1*** (1.8)	-3.3%	(-9.1, -3.1)	0.00	189	195	-3.0* (1.7)	-1.6%	(-5.8, -0.2)	0.08	0.22
PY 3	174	179	-6.9*** (2.1)	-3.8%	(-10.3, -3.5)	0.00	186	194	-4.7** (2.0)	-2.5%	(-8.0, -1.5)	0.02	0.45
PY 4	127	136	-10.3*** (2.4)	-7.5%	(-14.3, -6.4)	0.00	139	142	-1.0 (2.2)	-0.7%	(-4.6, 2.7)	0.66	0.00
PY 5	128	137	-10.5*** (2.4)	-7.6%	(-14.5, -6.5)	0.00	140	145	-2.7 (2.2)	-1.9%	(-6.3, 0.9)	0.22	0.02
PY 1 through 5	156	162	-7.8*** (1.8)	-4.8%	(-10.8, -4.9)	0.00	168	173	-3.0* (1.7)	-1.7%	(-5.7, -0.2)	0.07	0.05
Potentially primary care preventable outpatient ED visits[e]													
Baseline	127	125	NA	NA	NA	NA	137	136	NA	NA	NA	NA	NA
PY 1	124	125	-3.0*** (1.1)	-2.3%	(-4.8, -1.1)	0.01	134	135	-2.1* (1.1)	-1.6%	(-4.0, -0.3)	0.06	0.62
PY 2	122	122	-2.9** (1.3)	-2.3%	(-5.0, -0.8)	0.02	132	133	-1.8 (1.3)	-1.3%	(-4.0, 0.4)	0.18	0.56
PY 3	121	123	-3.5** (1.4)	-2.8%	(-5.8, -1.3)	0.01	132	133	-2.0 (1.3)	-1.5%	(-4.2, 0.2)	0.13	0.44
PY 4	92	96	-6.5*** (1.5)	-6.6%	(-9.0, -3.9)	0.00	102	101	0.4 (1.4)	0.4%	(-1.9, 2.7)	0.80	0.00
PY 5	97	101	-5.7*** (1.6)	-5.5%	(-8.2, -3.1)	0.00	106	106	-0.9 (1.4)	-0.8%	(-3.2, 1.5)	0.55	0.02
PY 1 through 5	110	112	-4.3*** (1.2)	-3.7%	(-6.2, -2.3)	0.00	120	121	-1.3 (1.1)	-1.1%	(-3.2, 0.5)	0.23	0.07

Table 5.A.1.2b. ***(continued)***

	Track 2 — SSP						Track 2 — Non-SSP						
	CPC+ mean[a]	C mean[a]	Impact estimate[b] (SE)	Percentage impact[c]	90% confidence interval	*p*-Value	CPC+ mean[a]	C mean[a]	Impact estimate[b] (SE)	Percentage impact[c]	90% confidence interval	*p*-Value	*p*-Value for SSP vs. non-SSP difference
Total Urgent Care Center (UCC) visits													
Baseline	99	104	NA	NA	NA	NA	96	107	NA	NA	NA	NA	NA
PY 1	115	116	3.8 (3.8)	3.4%	(-2.5, 10.0)	0.32	108	120	-0.9 (2.2)	-0.8%	(-4.6, 2.8)	0.69	0.29
PY 2	132	128	8.5* (4.6)	6.9%	(1.0, 16.1)	0.06	118	132	-2.8 (3.8)	-2.3%	(-8.9, 3.4)	0.46	0.06
PY 3	138	143	-0.1 (6.1)	-0.1%	(-10.2, 9.9)	0.98	131	146	-4.0 (4.3)	-3.0%	(-11.1, 3.1)	0.36	0.61
PY 4	141	136	10.2 (6.3)	7.8%	(-0.1, 20.6)	0.10	131	135	7.7 (4.8)	6.2%	(-0.1, 15.6)	0.11	0.76
PY 5	196	205	-4.2 (9.4)	-2.1%	(-19.7, 11.2)	0.65	179	180	9.8 (7.7)	5.8%	(-2.8, 22.4)	0.20	0.24
PY 1 through 5	146	147	3.6 (4.9)	2.5%	(-4.5, 11.7)	0.46	135	144	1.6 (3.9)	1.2%	(-4.7, 8.0)	0.67	0.76
Primary care substitutable UCC visits													
Baseline	59	62	NA	NA	NA	NA	57	63	NA	NA	NA	NA	NA
PY 1	69	70	2.6 (2.3)	3.9%	(-1.2, 6.3)	0.26	65	71	-1.2 (1.4)	-1.8%	(-3.4, 1.1)	0.38	0.16
PY 2	79	77	4.9* (2.7)	6.6%	(0.5, 9.2)	0.07	70	78	-2.5 (2.2)	-3.4%	(-6.1, 1.2)	0.27	0.04
PY 3	84	85	1.5 (3.7)	1.8%	(-4.5, 7.5)	0.68	78	86	-2.5 (2.5)	-3.0%	(-6.6, 1.7)	0.33	0.38
PY 4	93	87	8.7** (3.8)	10.3%	(2.4, 15.0)	0.02	84	84	5.4* (3.0)	6.9%	(0.6, 10.3)	0.07	0.53
PY 5	79	74	7.9** (3.9)	11.1%	(1.4, 14.4)	0.04	74	75	4.3 (3.3)	6.3%	(-1.1, 9.8)	0.19	0.49
PY 1 through 5	81	79	5.0* (2.7)	6.6%	(0.5, 9.5)	0.07	75	79	0.5 (2.1)	0.7%	(-2.9, 4.0)	0.80	0.20
UCC visits that excludes COVID-related diagnoses													
Baseline	98	103	NA	NA	NA	NA	96	107	NA	NA	NA	NA	NA
PY 1	115	116	3.4 (3.7)	3.0%	(-2.8, 9.5)	0.37	108	120	-0.9 (2.2)	-0.9%	(-4.6, 2.7)	0.68	0.32
PY 2	131	127	8.1* (4.6)	6.6%	(0.6, 15.6)	0.08	118	131	-2.4 (3.7)	-2.0%	(-8.6, 3.7)	0.52	0.08
PY 3	138	143	-0.3 (6.1)	-0.2%	(-10.3, 9.7)	0.96	130	145	-3.7 (4.3)	-2.7%	(-10.7, 3.4)	0.40	0.65
PY 4	99	100	3.7 (5.9)	3.8%	(-6.1, 13.4)	0.54	95	105	1.4 (4.2)	1.5%	(-5.5, 8.3)	0.74	0.74
PY 5	115	117	3.4 (6.1)	3.0%	(-6.7, 13.5)	0.58	114	124	0.6 (5.4)	0.5%	(-8.3, 9.4)	0.92	0.75

Table 5.A.1.2b. ***(continued)***

	Track 2 — SSP						Track 2 — Non-SSP						
	CPC+ mean[a]	C mean[a]	Impact estimate[b] (SE)	Percentage impact[c]	90% confidence interval	*p*-Value	CPC+ mean[a]	C mean[a]	Impact estimate[b] (SE)	Percentage impact[c]	90% confidence interval	*p*-Value	*p*-Value for SSP vs. non-SSP difference
PY 1 through 5	119	120	3.6 (4.5)	3.1%	(-3.8, 11.0)	0.42	113	125	-1.1 (3.5)	-1.0%	(-6.8, 4.5)	0.74	0.40
Ambulatory primary care visits (including to FQHCs, RHCs, and CAHs)[f]													
Baseline	4,214	4,355	NA	NA	NA	NA	4,476	4,504	NA	NA	NA	NA	NA
PY 1	4,237	4,425	-47.2** (20.9)	-1.1%	(-81.5, -12.8)	0.02	4,466	4,584	-90.4*** (23.9)	-2.0%	(-129.8, -51.1)	0.00	0.17
PY 2	4,268	4,436	-27.5 (28.7)	-0.6%	(-74.7, 19.7)	0.34	4,494	4,581	-59.8* (31.0)	-1.3%	(-110.8, -8.8)	0.05	0.44
PY 3	4,333	4,492	-18.7 (38.4)	-0.4%	(-81.9, 44.6)	0.63	4,542	4,617	-48.5 (36.1)	-1.1%	(-108.0, 11.0)	0.18	0.57
PY 4	3,913	4,055	-0.8 (39.4)	0.0%	(-65.7, 64.1)	0.98	4,103	4,182	-52.2 (36.8)	-1.3%	(-112.7, 8.3)	0.16	0.35
PY 5	4,161	4,311	-9.4 (48.1)	-0.2%	(-88.5, 69.7)	0.84	4,295	4,389	-66.3 (41.7)	-1.5%	(-134.9, 2.2)	0.11	0.38
PY 1 through 5	4,178	4,340	-20.8 (30.1)	-0.5%	(-70.3, 28.8)	0.49	4,373	4,463	-63.0** (30.0)	-1.4%	(-112.4, -13.7)	0.04	0.33
Ambulatory specialty care visits (including to FQHCs, RHCs, and CAHs)[f]													
Baseline	4,638	4,511	NA	NA	NA	NA	4,258	4,172	NA	NA	NA	NA	NA
PY 1	4,564	4,457	-19.3 (17.3)	-0.4%	(-47.8, 9.1)	0.26	4,233	4,136	10.6 (12.3)	0.3%	(-9.5, 30.8)	0.39	0.16
PY 2	4,534	4,462	-54.1** (24.0)	-1.2%	(-93.6, -14.6)	0.02	4,223	4,117	20.5 (17.3)	0.5%	(-8.0, 48.9)	0.24	0.01
PY 3	4,435	4,378	-69.9*** (26.1)	-1.6%	(-112.8, -26.9)	0.01	4,139	4,033	20.3 (21.0)	0.5%	(-14.3, 54.8)	0.33	0.01
PY 4	3,820	3,784	-90.4*** (28.5)	-2.3%	(-137.3, -43.4)	0.00	3,595	3,479	29.6 (23.6)	0.8%	(-9.3, 68.4)	0.21	0.00
PY 5	4,179	4,149	-96.6*** (30.0)	-2.3%	(-146.0, -47.3)	0.00	3,892	3,824	-18.3 (26.8)	-0.5%	(-62.4, 25.8)	0.49	0.05
PY 1 through 5	4,292	4,232	-66.0*** (22.0)	-1.5%	(-102.1, -29.9)	0.00	4,003	3,904	12.9 (17.7)	0.3%	(-16.2, 42.0)	0.47	0.01
Unweighted sample sizes for measures per 1,000 beneficiaries per year[g]													
Number of practices	636	1,817					879	1,966					
Number of beneficiaries	847,208	2,257,322					1,053,634	2,261,852					
Number of beneficiary-years	3,204,963	8,538,135					4,020,326	8,516,384					

Source: Mathematica's analysis of Medicare claims data from January 2013 through December 2021.

Notes: This table indicates which estimates are statistically significant; when we interpret evidence, we combine evidence from the magnitude of the effect, the *p*-values, findings on related outcomes, subgroups, sensitivity tests, and other data sources about model implementation.

Table 5.A.1.2b. ***(continued)***

[a] We report the actual, unadjusted averages in the baseline period which are similar for the CPC+ and comparison groups due to matching. In the intervention periods, the comparison group mean is computed by subtracting the regression adjusted difference between the CPC+ and comparison means in each time period from the CPC+ mean in that same time period.

[b] Each impact estimate is regression-adjusted using a difference-in-differences analysis that reflects the difference of the average outcome for Medicare FFS beneficiaries attributed to CPC+ practices in the five years of CPC+ to the average outcome in the baseline year, relative to the same difference over time for Medicare FFS beneficiaries attributed to comparison practices, while controlling for beneficiary characteristics and practice fixed effects.

[c] We calculated percentage impacts relative to what the CPC+ mean would have been in Program Years 1 through 5 (separately and combined) in the absence of the intervention—that is, the unadjusted CPC+ mean minus the impact estimate.

[d] Total ED visits include ED/observation stays that led to a hospitalization, including a psychiatric hospitalization.

[e] The sum of primary care substitutable outpatient ED visits and potentially primary care preventable outpatient ED visits is less than total outpatient ED visits because total outpatient ED visits include those for other care needs, such as injuries, mental health, drugs, and alcohol.

[f] Ambulatory visits with primary care practitioners and specialists include office-based visits and visits at home, as well as visits in other settings, such as FQHCs, RHCs, and CAHs.

[g] After accounting for weights that adjust for matching and time observed in Medicare FFS, the effective sample sizes fall but are still substantial. For the comparison group, the effective sample size is 38 to 43 percent of the size of the actual comparison group. The effective sample size for the CPC+ group is 96 percent of the actual sample size because it is affected only by time observed (and not by the matching weights).

*/**/*** Significantly different from zero at the 0.10/0.05/0.01 level, two-tailed test.

C = comparison; CAH = critical access hospital; ED = emergency department; FFS = fee-for-service; FQHC = federally qualified health center; NA = not applicable; PY = Program Year; RHC = rural health clinic; SE = standard error; SSP = Medicare Shared Savings Program.

Table 5.A.1.3a. Estimated average annual impacts of CPC+ on outpatient ED visits across the five program years, by baseline practice characteristics, Track 1

Practice subgroup definition, based on baseline characteristics	Track 1 – Overall			
	Number (percentage) of CPC+ beneficiaries in subgroup at baseline	Impact estimate (standard error)	Percentage impact	*p*-Value for difference in impact estimates between subgroups[a]
Main analysis (all practices)- average annual estimate for PY 1 through PY 5	-	**-9.4*** (2.5)**	**-2.1%**	-
Whether practice participated in prior primary care transformation initiatives (recognized as a medical home or participated in MAPCP or CPC Classic)				
Yes	468,487 (53.6%)	-13.0 (3.7)	-2.8%	
No	405,383 (46.4%)	-5.6 (3.3)	-1.2%	0.15
Large and medium versus small practice based on number of primary care practitioners				
Large (6+ primary care practitioners)	404,456 (46.3%)	-13.7 (3.9)	-3.0%	
Medium (3–5 primary care practitioners)	282,380 (32.3%)	-2.3 (4.1)	-0.5%	
Small (1–2 primary care practitioners)	187,034 (21.4%)	-11.4 (5.1)	-2.5%	0.15
Whether hospital- or system-owned versus independent (based on IQVIA data)[b]				
Hospital- or system-owned	474,606 (54.3%)	-7.2 (3.6)	-1.5%	
Independent	399,264 (45.7%)	-12.3 (3.3)	-2.9%	0.27
Whether the practice shared a TIN with another primary care practice[b]				
Shared a TIN with another primary care practice	684,507 (78.3%)	-8.4 (2.9)	-1.8%	
Did not share a TIN with another primary care practice	189,364 (21.7%)	-11.7 (4.6)	-2.7%	0.46
Practice type: multi-specialty versus primary care only				
Multi-specialty	170,691 (19.5%)	-9.6 (6.3)	-2.1%	
Primary care only	703,179 (80.5%)	-9.5 (2.7)	-2.1%	0.66
Urbanicity of practice's county: rural or suburban location versus urban location				
Rural	89,834 (10.3%)	-11.1 (9.2)	-2.0%	
Suburban	156,799 (17.9%)	0.1 (6.6)	0.0%	
Urban	627,237 (71.8%)	-11.7 (2.8)	-2.7%	0.39

Source: Mathematica's analysis of Medicare claims data from January 2013 through December 2021.

Note: The estimates (and standard errors) in the impact estimate column show subgroup*p*-specific impacts over the five years of CPC+, separately, for each practice characteristic listed in the table. We only tested differences within each subgroup if the estimates were significantly different between the two subgroups (that is, the *p*-value in the last column was <.10). Asterisks denote whether the impact estimate within a subgroup was significantly different from zero when estimates were significantly different between the subgroup categories.

[a] The *p*-values in the last column represent results from testing for statistically significant differences in impact estimates between the subgroups, based on the baseline practice characteristic (using a t-test for subgroups with two categories and from an F-test for subgroups with more than two categories).

[b] Since ownership status of a practice is likely to be highly correlated with whether the practice shares TIN with other practices, we included only one of these characteristics at a time in these regressions.

*/**/***Within-subgroup estimate significantly different from zero at the 0.10/0.05/0.01 level, two-tailed test.

CPC = Comprehensive Primary Care; ED = emergency department; MAPCP = Multi-payer Advanced Primary Care Practice Demonstration; PY = Program Year; TIN = Tax Identification Number.

Table 5.A.1.3b. Estimated average annual impacts of CPC+ on outpatient ED visits across the five program years, by baseline practice characteristics and SSP status, Track 1

Practice subgroup definition, based on baseline characteristics	Track 1 – SSP				Track 1 – Non-SSP			
	Number (percentage) of CPC+ beneficiaries in subgroup at baseline	Impact estimate (standard error)	Percentage impact	*p*-Value for difference in impact estimates between subgroups[a]	Number (percentage) of CPC+ beneficiaries in subgroup at baseline	Impact estimate (standard error)	Percentage impact	*p*-Value for difference in impact estimates between subgroups[a]
Main analysis (all practices)- average annual estimate for PY 1 through PY 5	**-**	**-10.2*** (3.3)**	**-2.3%**	**-**	**-**	**-8.2** (3.8)**	**-1.8%**	**-**
Whether practice participated in prior primary care transformation initiatives (recognized as a medical home or participated in MAPCP or CPC Classic)								
Yes	214,075 (47.7%)	-18.7*** (4.8)	-4.1%		254,262 (59.8%)	-7.6 (5.4)	-1.6%	
No	234,948 (52.3%)	-2.5 (4.4)	-0.6%	0.01	170,586 (40.2%)	-9.3 (5.0)	-2.0%	0.56
Large and medium versus small practice based on number of primary care practitioners								
Large (6+ primary care practitioners)	189,229 (42.1%)	-14.2 (4.9)	-3.2%		215,122 (50.6%)	-13.0 (5.9)	-2.8%	
Medium (3–5 primary care practitioners)	156,338 (34.8%)	-3.7 (5.7)	-0.8%		126,106 (29.7%)	0.3 (5.8)	0.1%	
Small (1–2 primary care practitioners)	103,455 (23.0%)	-12.9 (6.8)	-2.9%	0.39	83,621 (19.7%)	-9.1 (7.6)	-1.9%	0.23
Whether hospital- or system-owned versus independent (based on IQVIA data)[b]								
Hospital- or system-owned	250,558 (55.8%)	-5.8 (4.5)	-1.2%		224,086 (52.7%)	-8.3 (5.7)	-1.7%	
Independent	198,464 (44.2%)	-15.9 (4.6)	-3.8%	0.13	200,762 (47.3%)	-8.3 (4.9)	-1.9%	0.92
Whether the practice shared a TIN with another primary care practice[b]								
Shared a TIN with another primary care practice	366,843 (81.7%)	-7.0* (3.6)	-1.6%		317,749 (74.8%)	-9.3 (4.6)	-2.0%	
Did not share a TIN with another primary care practice	82,179 (18.3%)	-23.2*** (6.6)	-5.4%	0.03	107,099 (25.2%)	-3.7 (6.0)	-0.8%	0.52
Practice type: multi-specialty versus primary care only								
Multi-specialty	76,547 (17.0%)	-9.6 (7.8)	-2.3%		94,082 (22.1%)	-10.3 (9.4)	-2.2%	
Primary care only	372,475 (83.0%)	-10.4 (3.5)	-2.3%	0.45	330,766 (77.9%)	-7.7 (4.1)	-1.6%	0.93
Urbanicity of practice's county: rural or suburban location versus urban location								
Rural	22,327 (5.0%)	-17.6 (18.1)	-3.3%		67,372 (15.9%)	-10.5 (10.7)	-1.9%	
Suburban	74,982 (16.7%)	9.2 (10.1)	1.8%		81,785 (19.3%)	-7.3 (8.3)	-1.5%	
Urban	351,712 (78.3%)	-13.9 (3.3)	-3.3%	0.27	275,691 (64.9%)	-8.0 (4.6)	-1.8%	0.95

Source: Mathematica's analysis of Medicare claims data from January 2013 through December 2021.

Note: The estimates (and standard errors) in the impact estimate column show subgroup*p*-specific impacts over the five years of CPC+, separately, for each practice characteristic listed in the table. We only tested differences within each subgroup if the estimates were significantly different between the two subgroups (that is, the *p*-value in the last

Table 5.A.1.3b. ***(continued)***

column was <.10). Asterisks denote whether the impact estimate within a subgroup was significantly different from zero when estimates were significantly different between the subgroup categories.

[a] The *p*-values in the last column represent results from testing for statistically significant differences in impact estimates between the subgroups, based on the baseline practice characteristic (using a t-test for subgroups with two categories and from an F-test for subgroups with more than two categories).

[b] Since ownership status of a practice is likely to be highly correlated with whether the practice shares TIN with other practices, we included only one of these characteristics at a time in these regressions.

*/**/***Within-subgroup estimate significantly different from zero at the 0.10/0.05/0.01 level, two-tailed test.

CPC = Comprehensive Primary Care; ED = emergency department; MAPCP = Multi-payer Advanced Primary Care Practice Demonstration; PY = Program Year; SSP = Medicare Shared Savings Program; TIN = Tax Identification Number.

Table 5.A.1.4a. Estimated average annual impacts of CPC+ on outpatient ED visits across the five program years, by baseline practice characteristics, Track 2

Practice subgroup definition, based on baseline characteristics	Track 2 – Overall: Number (percentage) of CPC+ beneficiaries in subgroup at baseline	Track 2 – Overall: Impact estimate (standard error)	Track 2 – Overall: Percentage impact	Track 2 – Overall: *p*-Value for difference in impact estimates between subgroups[a]
Main analysis (all practices)- average annual estimate for PY 1 through PY 5	**-**	**-8.3*** (2.6)**	**-1.8%**	**-**
Whether practice participated in prior primary care transformation initiatives (recognized as a medical home or participated in MAPCP or CPC Classic)				
Yes	865,798 (81.2%)	-9.1 (3.1)	-2.0%	
No	201,028 (18.8%)	-6.4 (4.4)	-1.4%	0.70
Large and medium versus small practice based on number of primary care practitioners				
Large (6+ primary care practitioners)	589,224 (55.2%)	-9.5(3.8)	-2.1%	
Medium (3–5 primary care practitioners)	340,406 (31.9%)	-5.5 (4.1)	-1.2%	
Small (1–2 primary care practitioners)	137,196 (12.9%)	-12.1 (6.5)	-2.6%	0.47
Whether hospital- or system-owned versus independent (based on IQVIA data)[b]				
Hospital- or system-owned	619,957 (58.1%)	-8.7(3.6)	-1.8%	
Independent	446,869 (41.9%)	-8.4 (3.7)	-2.0%	0.91
Whether the practice shared a TIN with another primary care practice[b]				
Shared a TIN with another primary care practice	913,196 (85.6%)	-6.9(2.9)	-1.5%	
Did not share a TIN with another primary care practice	153,630 (14.4%)	-12.3 (5.4)	-2.9%	0.17
Practice type: multi-specialty versus primary care only				
Multi-specialty	278,801 (26.1%)	-11.9 (5.9)	-2.4%	
Primary care only	788,025 (73.9%)	-7.4 (2.8)	-1.7%	0.28
Urbanicity of practice's county: rural or suburban location versus urban location				
Rural	82,613 (7.7%)	2.5 (9.1)	0.5%	
Suburban	170,323 (16.0%)	9.0 (7.1)	1.9%	
Urban	813,890 (76.3%)	-13.4*** (2.9)	-3.0%	0.00

Source: Mathematica's analysis of Medicare claims data from January 2013 through December 2021.

Note: The estimates (and standard errors) in the impact estimate column show subgroup-specific impacts over the five years of CPC+, separately, for each practice characteristic listed in the table. We only tested differences within each subgroup if the estimates were significantly different between the two subgroups (that is, the *p*-value in the last column was <.10). Asterisks denote whether the impact estimate within a subgroup was significantly different from zero when estimates were significantly different between the subgroup categories.

[a] The *p*-values in the last column represent results from testing for statistically significant differences in impact estimates between the subgroups, based on the baseline practice characteristic (using a t-test for subgroups with two categories and from an F-test for subgroups with more than two categories).

[b] Since ownership status of a practice is likely to be highly correlated with whether the practice shares TIN with other practices, we included only one of these characteristics at a time in these regressions.

*/**/***Within-subgroup estimate significantly different from zero at the 0.10/0.05/0.01 level, two-tailed test.

CPC = Comprehensive Primary Care; ED = emergency department; MAPCP = Multi-payer Advanced Primary Care Practice Demonstration; PY = Program Year; TIN = Tax Identification Number.

Table 5.A.1.4b. Estimated average annual impacts of CPC+ on outpatient ED visits across the five program years, by baseline practice characteristics and SSP status, Track 2

Practice subgroup definition, based on baseline characteristics	Track 2 – SSP				Track 2 – Non-SSP			
	Number (percentage) of CPC+ beneficiaries in subgroup at baseline	Impact estimate (standard error)	Percentage impact	*p*-Value for difference in impact estimates between subgroups[a]	Number (percentage) of CPC+ beneficiaries in subgroup at baseline	Impact estimate (standard error)	Percentage impact	*p*-Value for difference in impact estimates between subgroups[a]
Main analysis (all practices)- average annual estimate for PY 1 through PY 5	**-**	**-12.7*** (3.8)**	**-2.9%**	**-**	**-**	**-3.8 (3.5)**	**-0.8%**	**-**
Whether practice participated in prior primary care transformation initiatives (recognized as a medical home or participated in MAPCP or CPC Classic)								
Yes	385,875 (81.8%)	-15.7 (4.2)	-3.6%		479,947 (80.6%)	-2.8 (4.2)	-0.6%	
No	85,762 (18.2%)	-2.6 (7.0)	-0.6%	0.19	115,242 (19.4%)	-7.2 (5.5)	-1.5%	0.50
Large and medium versus small practice based on number of primary care practitioners								
Large (6+ primary care practitioners)	279,067 (59.2%)	-11.8 (4.9)	-2.6%		310,301 (52.1%)	-6.5 (5.4)	-1.4%	
Medium (3–5 primary care practitioners)	134,103 (28.4%)	-18.5 (6.3)	-4.2%		206,177 (34.6%)	4.1 (5.5)	0.9%	
Small (1–2 primary care practitioners)	58,467 (12.4%)	-8.8 (10.2)	-2.0%	0.62	78,712 (13.2%)	-12.6 (8.3)	-2.7%	0.13
Whether hospital- or system-owned versus independent (based on IQVIA data)[b]								
Hospital- or system-owned	289,350 (61.4%)	-11.8 (4.9)	-2.6%		330,724 (55.6%)	-4.5 (4.9)	-0.9%	
Independent	182,287 (38.6%)	-15.8 (5.3)	-3.8%	0.45	264,465 (44.4%)	-2.6 (5.2)	-0.6%	0.76
Whether the practice shared a TIN with another primary care practice[b]								
Shared a TIN with another primary care practice	416,348 (88.3%)	-10.9*** (3.9)	-2.4%		496,945 (83.5%)	-2.4 (4.1)	-0.5%	
Did not share a TIN with another primary care practice	55,289 (11.7%)	-22.6** (10.0)	-5.5%	0.10	98,244 (16.5%)	-6.0 (6.1)	-1.4%	0.60
Practice type: multi-specialty versus primary care only								
Multi-specialty	116,601 (24.7%)	-14.0 (8.7)	-2.9%		162,149 (27.2%)	-9.3 (7.7)	-1.9%	
Primary care only	355,036 (75.3%)	-13.1 (3.9)	-3.1%	0.58	433,040 (72.8%)	-1.5 (4.0)	-0.3%	0.24
Urbanicity of practice's county: rural or suburban location versus urban location								
Rural	18,533 (3.9%)	-30.0 (20.9)	-5.8%		63,941 (10.7%)	12.0 (9.8)	2.4%	
Suburban	75,938 (16.1%)	13.4 (11.1)	2.8%		94,390 (15.9%)	6.0 (8.3)	1.2%	
Urban	377,166 (80.0%)	-17.9*** (3.8)	-4.2%	0.03	436,858 (73.4%)	-8.0* (4.2)	-1.8%	0.03

Source: Mathematica's analysis of Medicare claims data from January 2013 through December 2021.

Note: The estimates (and standard errors) in the impact estimate column show subgroup*p*-specific impacts over the five years of CPC+, separately, for each practice characteristic listed in the table. We only tested differences within each subgroup if the estimates were significantly different between the two subgroups (that is, the *p*-value in the last

Table 5.A.1.4b. ***(continued)***

column was <.10). Asterisks denote whether the impact estimate within a subgroup was significantly different from zero when estimates were significantly different between the subgroup categories.

[a] The *p*-values in the last column represent results from testing for statistically significant differences in impact estimates between the subgroups, based on the baseline practice characteristic (using a t-test for subgroups with two categories and from an F-test for subgroups with more than two categories).

[b] Since ownership status of a practice is likely to be highly correlated with whether the practice shares TIN with other practices, we included only one of these characteristics at a time in these regressions.

*/**/***Within-subgroup estimate significantly different from zero at the 0.10/0.05/0.01 level, two-tailed test.

CPC = Comprehensive Primary Care; ED = emergency department; MAPCP = Multi-payer Advanced Primary Care Practice Demonstration; PY = Program Year; SSP = Medicare Shared Savings Program; TIN = Tax Identification Number.

Table 5.A.1.5a. Estimated average annual impacts of CPC+ on outpatient ED visits across the five program years, by baseline beneficiary characteristics, Track 1

Beneficiary subgroup definition, based on baseline characteristics	Track 1 – Overall			
	Number (percentage) of CPC+ beneficiaries in subgroup at baseline	Impact estimate (standard error)	Percentage impact	*p*-Value for difference in impact estimates between subgroups[a]
Main analysis (all beneficiaries) – average annual estimate for PY 1 through PY 5	**-**	**-9.4*** (2.5)**	**-2.1%**	**-**
Patients in the highest quartile of the HCC score distribution				
Yes	203,811 (25.9%)	-11.7 (6.8)	-1.3%	
No	583,156 (74.1%)	-6.2 (2.1)	-1.7%	0.41
Patients in the highest decile of the HCC score distribution or who have dementia				
Yes	123,085 (15.6%)	-10.8 (9.3)	-1.1%	
No	663,882 (84.4%)	-7.0 (2.2)	-1.7%	0.68
Patients with anxiety/depression or substance use disorders				
Yes	120,562 (16.6%)	-14.9 (9.8)	-1.8%	
No	604,012 (83.4%)	-5.4 (2.0)	-1.4%	0.33
Patients with multiple chronic conditions (at least 2 of 12 frequently occurring chronic conditions[b]) and one or more hospitalizations[c]				
Yes	68,204 (8.7%)	-5.1 (14.6)	-0.5%	
No	718,763 (91.3%)	-7.8 (2.3)	-1.8%	0.85
Patients dually eligible for Medicare and Medicaid				
Yes	107,885 (12.6%)	-28.5** (11.1)	-2.8%	
No	746,776 (87.4%)	-8.1*** (2.0)	-2.1%	0.06

Source: Mathematica's analysis of Medicare claims data from January 2013 through December 2021.

Note: Beneficiary characteristics to determine subgroup membership are measured at the start of the year-long baseline period for baseline observations and at the start of Program Year 1 for observations in the intervention period (Program Years 1 through 5). The estimates (and standard errors) in the impact estimate column show subgroup-specific impacts, separately for each beneficiary characteristic listed in the table. We only tested differences *within* each subgroup if the estimates were significantly different *between* the two subgroups (that is, the *p*-value in the last column was <.10). Asterisks denote whether the impact estimate *within* a subgroup was significantly different from zero when estimates were significantly different between the subgroup categories. Because we could not observe diagnoses (which are used to determine HCCs and calculate HCC scores) at baseline for beneficiaries who were new to Medicare during the program years, we excluded new Medicare beneficiaries from all subgroup analyses (except the analysis based on dual status since beneficiaries who are new to Medicare, by definition, could not have been enrolled in both Medicare and Medicaid prior to joining Medicare). Due to this process, about 20 percent of observations from the regressions were excluded for the subgroups defined by HCC score and chronic conditions. Therefore, the main impact estimate of -9.4 for Track 1 overall may not lie between the impact estimates for these subgroups.

[a] The *p*-values in the last column represent results from testing for statistically significant differences in impact estimates between the subgroups, based on the baseline beneficiary characteristic (using a t-test for all subgroups).

[b] The 12 frequently occurring chronic conditions are congestive heart failure, chronic obstructive pulmonary disease, history of acute myocardial infarction, ischemic heart disease, diabetes, metastatic cancer and acute leukemia, history of stroke, depression, dementia, atrial fibrillation, rheumatoid arthritis or osteoarthritis, and chronic kidney disease.

[c] For observations in the baseline year, hospitalizations are measured in 2015, the year before the start of the baseline year. For observations in the intervention period, hospitalizations are measured in 2016, the year before the start of Program Year 1.

*/**/*** Significantly different from zero at the 0.10/0.05/0.01 level, two-tailed test.

ED = emergency department; HCC = hierarchical condition category; PY = Program Year.

Table 5.A.1.5b. Estimated average annual impacts of CPC+ on outpatient ED visits across the five program years, by baseline beneficiary characteristics and SSP status, Track 1

Beneficiary subgroup definition, based on baseline characteristics	Track 1 – SSP				Track 1 – Non-SSP			
	Number (percentage) of CPC+ beneficiaries in subgroup at baseline	Impact estimate (standard error)	Percentage impact	*p*-Value for difference in impact estimates between subgroups[a]	Number (percentage) of CPC+ beneficiaries in subgroup at baseline	Impact estimate (standard error)	Percentage impact	*p*-Value for difference in impact estimates between subgroups[a]
Main analysis (all beneficiaries) – average annual estimate for PY 1 through PY 5	-	**-10.2*** (3.3)**	**-2.3%**	-	-	**-8.2** (3.8)**	**-1.8%**	-
Patients in the highest quartile of the HCC score distribution								
Yes	115,215 (26.8%)	-14.3 (8.5)	-1.7%		88,864 (25.0%)	-8.6 (10.7)	-0.9%	
No	315,425 (73.2%)	-7.8 (2.7)	-2.2%	0.45	266,666 (75.0%)	-4.1 (3.1)	-1.1%	0.67
Patients in the highest decile of the HCC score distribution or who have dementia								
Yes	68,759 (16.0%)	-16.1 (12.0)	-1.7%		54,382 (15.3%)	-4.7 (14.4)	-0.5%	
No	361,881 (84.0%)	-8.1 (2.9)	-2.1%	0.50	301,148 (84.7%)	-5.5 (3.4)	-1.3%	0.95
Patients with anxiety/depression or substance use disorders								
Yes	66,746 (16.8%)	-17.3 (13.2)	-2.2%		53,792 (16.4%)	-11.4 (14.6)	-1.3%	
No	329,703 (83.2%)	-5.6 (2.5)	-1.5%	0.37	273,568 (83.6%)	-4.8 (3.0)	-1.2%	0.65
Patients with multiple chronic conditions (at least 2 of 12 frequently occurring chronic conditions[b]) and one or more hospitalizations[c]								
Yes	38,153 (8.9%)	-11.0 (18.2)	-1.0%		30,089 (8.5%)	2.2 (23.0)	0.2%	
No	392,487 (91.1%)	-9.4 (3.1)	-2.3%	0.93	325,442 (91.5%)	-5.9 (3.4)	-1.3%	0.72
Patients dually eligible for Medicare and Medicaid								
Yes	55,728 (11.9%)	-35.3** (15.1)	-3.6%		51,626 (13.3%)	-20.7 (16.2)	-2.0%	
No	410,653 (88.1%)	-8.8*** (2.6)	-2.3%	0.08	335,619 (86.7%)	-7.0 (3.1)	-1.8%	0.40

Source: Mathematica's analysis of Medicare claims data from January 2013 through December 2021.

Note: Beneficiary characteristics to determine subgroup membership are measured at the start of the year-long baseline period for baseline observations and at the start of Program Year 1 for observations in the intervention period (Program Years 1 through 5). The estimates (and standard errors) in the impact estimate column show subgroup*p*-specific impacts, separately for each beneficiary characteristic listed in the table. We only tested differences *within* each subgroup if the estimates were significantly different *between* the two subgroups (that is, the *p*-value in the last column was <.10). Asterisks denote whether the impact estimate *within* a subgroup was significantly different from zero when estimates were significantly different between the subgroup categories. Because we could not observe diagnoses (which are used to determine HCCs and calculate HCC scores) at baseline for beneficiaries who were new to Medicare during the program years, we excluded new Medicare beneficiaries from all subgroup analyses (except the analysis based on dual status since beneficiaries who are new to Medicare, by definition, could not have been enrolled in both Medicare and Medicaid prior to joining Medicare). Due to this process, about 20 percent of observations from the regressions were excluded for the subgroups defined by HCC score and chronic conditions. Therefore, the main impact estimate of -10.2 for Track 1 SSP and -8.2 for Track 1 Non-SSP may not lie between the impact estimates for these subgroups.

[a] The *p*-values in the last column represent results from testing for statistically significant differences in impact estimates between the subgroups, based on the baseline beneficiary characteristic (using a t-test for all subgroups).

Table 5.A.1.5b. ***(continued)***

[b] The 12 frequently occurring chronic conditions are congestive heart failure, chronic obstructive pulmonary disease, history of acute myocardial infarction, ischemic heart disease, diabetes, metastatic cancer and acute leukemia, history of stroke, depression, dementia, atrial fibrillation, rheumatoid arthritis or osteoarthritis, and chronic kidney disease.

[c] For observations in the baseline year, hospitalizations are measured in 2015, the year before the start of the baseline year. For observations in the intervention period, hospitalizations are measured in 2016, the year before the start of Program Year 1.

*/**/*** Significantly different from zero at the 0.10/0.05/0.01 level, two-tailed test.

ED = emergency department; HCC = hierarchical condition category; PY = Program Year; SSP = Medicare Shared Savings Program.

Table 5.A.1.6a. Estimated average annual impacts of CPC+ on outpatient ED visits across the five program years, by baseline beneficiary characteristics, Track 2

Beneficiary subgroup definition, based on baseline characteristics	Track 2 - Overall			
	Number (percentage) of CPC+ beneficiaries in subgroup at baseline	Impact estimate (standard error)	Percentage impact	*p*-Value for difference in impact estimates between subgroups[a]
Main analysis (all beneficiaries) – average annual estimate for PY 1 through PY 5	-	**-8.3*** (2.6)**	**-1.8%**	-
Patients in the highest quartile of the HCC score distribution				
Yes	268,430 (26.1%)	0.0 (7.1)	0.0%	
No	761,970 (73.9%)	-8.0 (2.2)	-2.2%	0.26
Patients in the highest decile of the HCC score distribution or who have dementia				
Yes	162,510 (15.8%)	-0.8 (9.4)	-0.1%	
No	867,891 (84.2%)	-6.9 (2.4)	-1.7%	0.52
Patients with anxiety/depression or substance use disorders				
Yes	164,048 (17.3%)	-17.2 (10.2)	-2.1%	
No	784,877 (82.7%)	-3.6 (2.0)	-1.0%	0.18
Patients with multiple chronic conditions (at least 2 of 12 frequently occurring chronic conditions[b]) and one or more hospitalizations[c]				
Yes	90,543 (8.8%)	3.3 (14.7)	0.3%	
No	939,858 (91.2%)	-6.8 (2.4)	-1.6%	0.49
Patients dually eligible for Medicare and Medicaid				
Yes	140,782 (12.5%)	-13.5 (11.8)	-1.4%	
No	984,688 (87.5%)	-8.6 (2.2)	-2.2%	0.68

Source: Mathematica's analysis of Medicare claims data from January 2013 through December 2021.

Note: Beneficiary characteristics to determine subgroup membership are measured at the start of the year-long baseline period for baseline observations and at the start of Program Year 1 for observations in the intervention period (Program Years 1 through 5). The estimates (and standard errors) in the impact estimate column show subgroup-specific impacts, separately for each beneficiary characteristic listed in the table. We only tested differences *within* each subgroup if the estimates were significantly different *between* the two subgroups (that is, the *p*-value in the last column was <.10). Asterisks denote whether the impact estimate *within* a subgroup was significantly different from zero when estimates were significantly different between the subgroup categories. Because we could not observe diagnoses (which are used to determine HCCs and calculate HCC scores) at baseline for beneficiaries who were new to Medicare during the program years, we excluded new Medicare beneficiaries from all subgroup analyses (except the analysis based on dual status since beneficiaries who are new to Medicare, by definition, could not have been enrolled in both Medicare and Medicaid prior to joining Medicare). Due to this process, about 20 percent of observations from the regressions were excluded for the subgroups defined by HCC score and chronic conditions. Therefore, the main impact estimate of -8.3 for Track 2 overall may not lie between the impact estimates for these subgroups.

[a] The *p*-values in the last column represent results from testing for statistically significant differences in impact estimates between the subgroups, based on the baseline beneficiary characteristic (using a t-test for all subgroups).

[b] The 12 frequently occurring chronic conditions are congestive heart failure, chronic obstructive pulmonary disease, history of acute myocardial infarction, ischemic heart disease, diabetes, metastatic cancer and acute leukemia, history of stroke, depression, dementia, atrial fibrillation, rheumatoid arthritis or osteoarthritis, and chronic kidney disease.

[c] For observations in the baseline year, hospitalizations are measured in 2015, the year before the start of the baseline year. For observations in the intervention period, hospitalizations are measured in 2016, the year before the start of Program Year 1.

*/**/*** Significantly different from zero at the 0.10/0.05/0.01 level, two-tailed test.

ED = emergency department; HCC = hierarchical condition category; PY = Program Year.

Table 5.A.1.6b. Estimated average annual impacts of CPC+ on outpatient ED visits across the five program years, by baseline beneficiary characteristics and SSP status, Track 2

Beneficiary subgroup definition, based on baseline characteristics	Track 2 – SSP				Track 2 – Non-SSP			
	Number (percentage) of CPC+ beneficiaries in subgroup at baseline	Impact estimate (standard error)	Percentage impact	*p*-Value for difference in impact estimates between subgroups[a]	Number (percentage) of CPC+ beneficiaries in subgroup at baseline	Impact estimate (standard error)	Percentage impact	*p*-Value for difference in impact estimates between subgroups[a]
Main analysis (all beneficiaries) – average annual estimate for PY 1 through PY 5	-	**-12.7*** (3.8)**	**-2.9%**	-	-	**-3.8 (3.5)**	**-0.8%**	-
Patients in the highest quartile of the HCC score distribution								
Yes	120,947 (26.8%)	-16.4 (9.8)	-2.0%		146,522 (25.5%)	14.4 (9.9)	1.6%	
No	330,277 (73.2%)	-7.9 (3.1)	-2.3%	0.38	428,947 (74.5%)	-7.4** (2.9)	-2.0%	0.03
Patients in the highest decile of the HCC score distribution or who have dementia								
Yes	71,030 (15.7%)	-19.2 (13.9)	-2.0%		90,895 (15.8%)	14.7 (12.7)	1.5%	
No	380,194 (84.3%)	-8.4 (3.4)	-2.2%	0.44	484,574 (84.2%)	-5.0 (3.3)	-1.2%	0.13
Patients with anxiety/depression or substance use disorders								
Yes	74,382 (17.8%)	-23.4 (14.0)	-3.0%		89,058 (16.8%)	-12.2 (14.5)	-1.4%	
No	342,453 (82.2%)	-6.8 (3.0)	-1.9%	0.23	439,501 (83.2%)	-0.6 (2.8)	-0.2%	0.42
Patients with multiple chronic conditions (at least 2 of 12 frequently occurring chronic conditions[b]) and one or more hospitalizations[c]								
Yes	41,080 (9.1%)	-40.3 (20.2)	-3.7%		49,139 (8.5%)	40.2* (20.9)	3.6%	
No	410,144 (90.9%)	-7.2 (3.6)	-1.7%	0.10	526,331 (91.5%)	-5.8* (3.1)	-1.3%	0.03
Patients dually eligible for Medicare and Medicaid								
Yes	55,837 (11.3%)	-32.6 (17.6)	-3.3%		84,414 (13.5%)	1.2 (15.6)	0.1%	
No	438,154 (88.7%)	-11.4 (3.3)	-2.9%	0.22	542,895 (86.5%)	-5.5 (2.8)	-1.4%	0.67

Source: Mathematica's analysis of Medicare claims data from January 2013 through December 2021.

Note: Beneficiary characteristics to determine subgroup membership are measured at the start of the year-long baseline period for baseline observations and at the start of Program Year 1 for observations in the intervention period (Program Years 1 through 5). The estimates (and standard errors) in the impact estimate column show subgroup-specific impacts, separately for each beneficiary characteristic listed in the table. We only tested differences *within* each subgroup if the estimates were significantly different *between* the two subgroups (that is, the *p*-value in the last column was <.10). Asterisks denote whether the impact estimate *within* a subgroup was significantly different from zero when estimates were significantly different between the subgroup categories. Because we could not observe diagnoses (which are used to determine HCCs and calculate HCC scores) at baseline for beneficiaries who were new to Medicare during the program years, we excluded new Medicare beneficiaries from all subgroup analyses (except the analysis based on dual status since beneficiaries who are new to Medicare, by definition, could not have been enrolled in both Medicare and Medicaid prior to joining Medicare). Due to this process, about 20 percent of observations from the regressions were excluded for the subgroups defined by HCC score and chronic conditions. Therefore, the main impact estimate of -12.7 for Track 2 SSP and -3.8 for Track 2 Non-SSP may not lie between the impact estimates for these subgroups.

[a] The p-values in the last column represent results from testing for statistically significant differences in impact estimates between the subgroups, based on the baseline beneficiary characteristic (using a t-test for all subgroups).

Table 5.A.1.6b. ***(continued)***

[b] The 12 frequently occurring chronic conditions are congestive heart failure, chronic obstructive pulmonary disease, history of acute myocardial infarction, ischemic heart disease, diabetes, metastatic cancer and acute leukemia, history of stroke, depression, dementia, atrial fibrillation, rheumatoid arthritis or osteoarthritis, and chronic kidney disease.

[c] For observations in the baseline year, hospitalizations are measured in 2015, the year before the start of the baseline year. For observations in the intervention period, hospitalizations are measured in 2016, the year before the start of Program Year 1.

*/**/*** Significantly different from zero at the 0.10/0.05/0.01 level, two-tailed test.

ED = emergency department; HCC = hierarchical condition category; PY = Program Year; SSP = Medicare Shared Savings Program.

Table 5.A.1.7a. Regression-adjusted means and estimated impacts of CPC+ on types of acute hospitalizations for attributed Medicare FFS beneficiaries by program year and average across the five program years, Track 1

	Track 1 — Overall					
	CPC+ mean[a]	C mean[a]	Impact estimate[b] (SE)	Percentage impact[c]	90% confidence interval	*p*-Value
Hospitalizations per 1,000 beneficiaries per year (short-stay acute care and critical access hospitals)						
Acute surgical hospitalizations						
Baseline	90	88	NA	NA	NA	NA
PY 1	89	85	1.0 (0.7)	1.2%	(-0.1, 2.1)	0.12
PY 2	86	84	0.1 (0.7)	0.2%	(-0.9, 1.2)	0.82
PY 3	87	85	-0.1 (0.7)	-0.1%	(-1.2, 1.1)	0.92
PY 4	71	70	-0.9 (0.7)	-1.2%	(-2.0, 0.2)	0.18
PY 5	69	67	-0.6 (0.7)	-0.9%	(-1.8, 0.5)	0.35
PY 1 through 5	80	78	-0.1 (0.5)	-0.1%	(-1.0, 0.8)	0.88
Acute surgical hospitalizations with a major complication or comorbidity						
Baseline	22	22	NA	NA	NA	NA
PY 1	22	21	0.5 (0.3)	2.2%	(-0.1, 1.0)	0.14
PY 2	22	22	0.0 (0.3)	-0.2%	(-0.5, 0.5)	0.91
PY 3	22	22	0.1 (0.3)	0.4%	(-0.4, 0.6)	0.77
PY 4	21	20	0.0 (0.3)	-0.2%	(-0.6, 0.5)	0.92
PY 5	21	21	-0.3 (0.3)	-1.2%	(-0.8, 0.3)	0.44
PY 1 through 5	22	21	0.0 (0.3)	0.2%	(-0.4, 0.5)	0.86
Acute surgical hospitalizations with a complication or comorbidity						
Baseline	22	22	NA	NA	NA	NA
PY 1	21	21	0.3 (0.3)	1.4%	(-0.2, 0.8)	0.33
PY 2	22	21	0.2 (0.3)	0.9%	(-0.3, 0.7)	0.54
PY 3	22	22	-0.1 (0.3)	-0.2%	(-0.6, 0.5)	0.87
PY 4	20	19	0.3 (0.3)	1.7%	(-0.2, 0.8)	0.28
PY 5	20	20	0.1 (0.3)	0.6%	(-0.4, 0.6)	0.72
PY 1 through 5	21	20	0.2 (0.2)	0.8%	(-0.2, 0.6)	0.50
Acute surgical hospitalizations without any complication or comorbidity						
Baseline	46	44	NA	NA	NA	NA
PY 1	45	44	0.3 (0.5)	0.6%	(-0.5, 1.0)	0.57
PY 2	43	41	0.0 (0.5)	0.0%	(-0.8, 0.7)	0.98
PY 3	42	41	-0.1 (0.4)	-0.3%	(-0.8, 0.6)	0.80
PY 4	31	30	-1.2*** (0.4)	-3.7%	(-1.9, -0.4)	0.01
PY 5	28	27	-0.5 (0.4)	-1.8%	(-1.2, 0.2)	0.27

Table 5.A.1.7a. ***(continued)***

	Track 1 — Overall					
	CPC+ mean[a]	C mean[a]	Impact estimate[b] (SE)	Percentage impact[c]	90% confidence interval	*p*-Value
PY 1 through 5	37	36	-0.3 (0.4)	-0.8%	(-0.9, 0.3)	0.41
Acute medical hospitalizations						
Baseline	200	201	NA	NA	NA	NA
PY 1	200	203	-1.7 (1.2)	-0.8%	(-3.7, 0.4)	0.18
PY 2	198	201	-2.0 (1.4)	-1.0%	(-4.3, 0.3)	0.15
PY 3	197	201	-2.6* (1.5)	-1.3%	(-5.0, -0.1)	0.08
PY 4	172	177	-4.0*** (1.5)	-2.3%	(-6.5, -1.5)	0.01
PY 5	175	178	-2.0 (1.5)	-1.1%	(-4.5, 0.6)	0.20
PY 1 through 5	188	191	-2.4** (1.2)	-1.3%	(-4.4, -0.4)	0.04
Acute medical hospitalizations with a major complication or comorbidity						
Baseline	75	75	NA	NA	NA	NA
PY 1	82	83	-0.2 (0.7)	-0.2%	(-1.4, 1.0)	0.81
PY 2	85	85	0.1 (0.8)	0.2%	(-1.2, 1.5)	0.86
PY 3	86	86	-0.6 (0.8)	-0.7%	(-2.0, 0.7)	0.44
PY 4	82	85	-2.3*** (0.9)	-2.7%	(-3.8, -0.9)	0.01
PY 5	87	87	0.2 (0.9)	0.2%	(-1.3, 1.7)	0.83
PY 1 through 5	84	85	-0.5 (0.7)	-0.6%	(-1.7, 0.6)	0.42
Acute medical hospitalizations with a complication or comorbidity						
Baseline	48	49	NA	NA	NA	NA
PY 1	45	46	-0.4 (0.5)	-0.9%	(-1.3, 0.5)	0.46
PY 2	45	46	-0.3 (0.5)	-0.7%	(-1.2, 0.6)	0.56
PY 3	44	45	0.2 (0.6)	0.4%	(-0.7, 1.1)	0.72
PY 4	36	37	-0.3 (0.6)	-0.9%	(-1.3, 0.6)	0.54
PY 5	36	37	-0.4 (0.5)	-1.2%	(-1.3, 0.5)	0.42
PY 1 through 5	41	42	-0.2 (0.5)	-0.6%	(-1.0, 0.5)	0.59
Acute medical hospitalizations without any complication or comorbidity						
Baseline	77	77	NA	NA	NA	NA
PY 1	73	74	-1.1* (0.7)	-1.5%	(-2.2, 0.0)	0.10
PY 2	69	71	-1.8** (0.7)	-2.6%	(-3.0, -0.6)	0.01
PY 3	67	70	-2.1*** (0.8)	-3.0%	(-3.4, -0.8)	0.01
PY 4	53	55	-1.4* (0.8)	-2.5%	(-2.7, 0.0)	0.10
PY 5	53	55	-1.7** (0.8)	-3.2%	(-3.0, -0.4)	0.03
PY 1 through 5	62	64	-1.6*** (0.6)	-2.6%	(-2.7, -0.6)	0.01

Table 5.A.1.7a. ***(continued)***

	Track 1 — Overall					
	CPC+ mean[a]	C mean[a]	Impact estimate[b] (SE)	Percentage impact[c]	90% confidence interval	*p*-Value
Unweighted sample sizes for measures per 1,000 beneficiaries per year[d]						
Number of practices	1,373	5,243				
Number of beneficiaries	1,549,585	5,347,499				
Number of beneficiary-years	5,916,394	20,150,090				

Source: Mathematica's analysis of Medicare claims data from January 2013 through December 2021.

Notes: This table indicates which estimates are statistically significant; when we interpret evidence, we combine evidence from the magnitude of the effect, the *p*-values, findings on related outcomes, subgroups, sensitivity tests, and other data sources about model implementation.

[a] We report the actual, unadjusted averages in the baseline period which are similar for the CPC+ and comparison groups due to matching. In the intervention periods, the comparison group mean is computed by subtracting the regression adjusted difference between the CPC+ and comparison means in each time period from the CPC+ mean in that same time period.

[b] Each impact estimate is regression-adjusted using a difference-in-differences analysis that reflects the difference of the average outcome for Medicare FFS beneficiaries attributed to CPC+ practices in the five years of CPC+ to the average outcome in the baseline year, relative to the same difference over time for Medicare FFS beneficiaries attributed to comparison practices, while controlling for beneficiary characteristics and practice fixed effects.

[c] We calculated percentage impacts relative to what the CPC+ mean would have been in Program Years 1 through 5 (separately and combined) in the absence of the intervention—that is, the unadjusted CPC+ mean minus the impact estimate.

[d] After accounting for weights that adjust for matching and time observed in Medicare FFS, the effective sample sizes fall but are still substantial. For the comparison group, the effective sample size is 45 percent of the size of the actual comparison group. The effective sample size for the CPC+ group is 96 percent of the actual sample size because it is affected only by time observed (and not by the matching weights).

*/**/*** Significantly different from zero at the 0.10/0.05/0.01 level, two-tailed test.

C = comparison; FFS = fee-for-service; NA = not applicable; PY = Program Year; SE = standard error.

Table 5.A.1.7b. Regression-adjusted means and estimated impacts of CPC+ on types of acute hospitalizations for attributed Medicare FFS beneficiaries by program year and average across the five program years, Track 1 by SSP status

	Track 1 — SSP						Track 1 — Non-SSP						
	CPC+ mean[a]	C mean[a]	Impact estimate[b] (SE)	Percentage impact[c]	90% confidence interval	*p*-Value	CPC+ mean[a]	C mean[a]	Impact estimate[b] (SE)	Percentage impact[c]	90% confidence interval	*p*-Value	*p*-Value for SSP vs. non-SSP difference
Hospitalizations per 1,000 beneficiaries per year (short-stay acute care and critical access hospitals)													
Acute surgical hospitalizations													
Baseline	90	88	NA	NA	NA	NA	90	87	NA	NA	NA	NA	NA
PY 1	89	86	0.1 (0.8)	0.1%	(-1.3, 1.5)	0.91	89	84	2.0* (1.1)	2.4%	(0.3, 3.8)	0.05	0.15
PY 2	87	84	0.2 (0.9)	0.2%	(-1.2, 1.6)	0.85	86	83	0.1 (1.0)	0.1%	(-1.6, 1.8)	0.91	0.98
PY 3	87	85	-0.6 (0.9)	-0.6%	(-2.0, 0.9)	0.53	87	84	0.5 (1.1)	0.6%	(-1.3, 2.2)	0.65	0.45
PY 4	71	71	-2.0** (0.8)	-2.7%	(-3.4, -0.6)	0.02	72	69	0.3 (1.0)	0.4%	(-1.4, 2.0)	0.81	0.10
PY 5	70	70	-2.3** (0.9)	-3.2%	(-3.8, -0.8)	0.01	69	65	0.8 (1.0)	1.2%	(-0.9, 2.6)	0.42	0.02
PY 1 through 5	80	79	-0.9 (0.7)	-1.0%	(-2.0, 0.3)	0.22	80	77	0.7 (0.9)	0.9%	(-0.7, 2.1)	0.40	0.15
Acute surgical hospitalizations with a major complication or comorbidity													
Baseline	22	22	NA	NA	NA	NA	22	22	NA	NA	NA	NA	NA
PY 1	22	21	0.4 (0.4)	1.9%	(-0.3, 1.1)	0.33	22	21	0.5 (0.5)	2.5%	(-0.3, 1.3)	0.28	0.86
PY 2	22	22	0.2 (0.4)	1.0%	(-0.4, 0.9)	0.59	22	22	-0.3 (0.5)	-1.4%	(-1.1, 0.5)	0.51	0.39
PY 3	22	22	-0.3 (0.4)	-1.2%	(-1.0, 0.4)	0.51	23	21	0.5 (0.5)	2.3%	(-0.3, 1.3)	0.31	0.23
PY 4	21	21	-0.6 (0.4)	-2.8%	(-1.3, 0.1)	0.16	21	20	0.6 (0.5)	2.7%	(-0.3, 1.4)	0.27	0.08
PY 5	22	22	-0.7 (0.4)	-3.2%	(-1.4, 0.0)	0.11	21	20	0.1 (0.5)	0.4%	(-0.8, 0.9)	0.88	0.24
PY 1 through 5	22	22	-0.2 (0.3)	-0.8%	(-0.7, 0.4)	0.60	22	21	0.3 (0.4)	1.2%	(-0.4, 0.9)	0.52	0.41

Table 5.A.1.7b. ***(continued)***

	Track 1 — SSP						Track 1 — Non-SSP						
	CPC+ mean[a]	C mean[a]	Impact estimate[b] (SE)	Percentage impact[c]	90% confidence interval	*p*-Value	CPC+ mean[a]	C mean[a]	Impact estimate[b] (SE)	Percentage impact[c]	90% confidence interval	*p*-Value	*p*-Value for SSP vs. non-SSP difference
Acute surgical hospitalizations with a complication or comorbidity													
Baseline	22	22	NA	NA	NA	NA	22	22	NA	NA	NA	NA	NA
PY 1	21	21	-0.1 (0.4)	-0.6%	(-0.8, 0.5)	0.74	21	20	0.8* (0.5)	3.7%	(0.0, 1.5)	0.10	0.14
PY 2	22	21	0.4 (0.4)	1.7%	(-0.3, 1.1)	0.39	21	21	0.0 (0.5)	0.0%	(-0.8, 0.7)	0.99	0.54
PY 3	23	22	-0.1 (0.4)	-0.3%	(-0.7, 0.6)	0.89	22	22	0.0 (0.5)	-0.2%	(-0.8, 0.7)	0.93	0.98
PY 4	20	20	0.2 (0.4)	0.9%	(-0.5, 0.9)	0.66	20	19	0.5 (0.5)	2.8%	(-0.2, 1.3)	0.24	0.57
PY 5	20	20	-0.1 (0.4)	-0.6%	(-0.9, 0.6)	0.77	20	19	0.4 (0.5)	1.9%	(-0.4, 1.1)	0.44	0.45
PY 1 through 5	21	21	0.1 (0.3)	0.3%	(-0.5, 0.6)	0.87	21	20	0.3 (0.4)	1.5%	(-0.3, 0.9)	0.41	0.62
Acute surgical hospitalizations without any complication or comorbidity													
Baseline	46	44	NA	NA	NA	NA	46	44	NA	NA	NA	NA	NA
PY 1	45	44	-0.2 (0.6)	-0.4%	(-1.1, 0.8)	0.76	45	43	0.8 (0.7)	1.7%	(-0.4, 1.9)	0.30	0.32
PY 2	42	41	-0.4 (0.6)	-1.0%	(-1.4, 0.5)	0.45	43	41	0.4 (0.7)	1.1%	(-0.7, 1.6)	0.53	0.33
PY 3	42	41	-0.2 (0.6)	-0.5%	(-1.2, 0.7)	0.70	43	41	0.0 (0.7)	0.1%	(-1.1, 1.1)	0.97	0.78
PY 4	30	30	-1.6*** (0.6)	-4.9%	(-2.5, -0.6)	0.01	31	30	-0.8 (0.7)	-2.7%	(-2.0, 0.3)	0.21	0.43
PY 5	28	28	-1.5** (0.6)	-5.0%	(-2.4, -0.5)	0.01	28	26	0.4 (0.7)	1.5%	(-0.7, 1.5)	0.56	0.04
PY 1 through 5	37	36	-0.7 (0.5)	-1.9%	(-1.5, 0.0)	0.12	38	36	0.2 (0.6)	0.4%	(-0.7, 1.1)	0.77	0.22
Acute medical hospitalizations													
Baseline	200	201	NA	NA	NA	NA	199	201	NA	NA	NA	NA	NA
PY 1	200	204	-2.8* (1.6)	-1.4%	(-5.4, -0.1)	0.09	200	202	-0.5 (1.9)	-0.2%	(-3.6, 2.7)	0.81	0.36
PY 2	199	203	-2.4 (1.8)	-1.2%	(-5.4, 0.6)	0.19	197	200	-1.6 (2.1)	-0.8%	(-5.0, 1.9)	0.46	0.77
PY 3	199	204	-4.3** (1.9)	-2.1%	(-7.5, -1.2)	0.02	196	198	-0.5 (2.3)	-0.2%	(-4.2, 3.3)	0.83	0.19
PY 4	174	181	-6.1*** (2.0)	-3.4%	(-9.4, -2.7)	0.00	170	173	-1.6 (2.3)	-0.9%	(-5.4, 2.2)	0.48	0.15
PY 5	180	183	-2.8 (2.1)	-1.5%	(-6.3, 0.7)	0.19	171	172	-0.2 (2.2)	-0.1%	(-3.8, 3.5)	0.93	0.39
PY 1 through 5	190	194	-3.7** (1.6)	-1.9%	(-6.3, -1.1)	0.02	186	188	-0.9 (1.8)	-0.5%	(-3.9, 2.2)	0.64	0.25

Table 5.A.1.7b. ***(continued)***

	Track 1 — SSP						Track 1 — Non-SSP						
	CPC+ mean[a]	C mean[a]	Impact estimate[b] (SE)	Percentage impact[c]	90% confidence interval	*p*-Value	CPC+ mean[a]	C mean[a]	Impact estimate[b] (SE)	Percentage impact[c]	90% confidence interval	*p*-Value	*p*-Value for SSP vs. non-SSP difference
Acute medical hospitalizations with a major complication or comorbidity													
Baseline	75	75	NA	NA	NA	NA	74	75	NA	NA	NA	NA	NA
PY 1	82	83	-1.4 (0.9)	-1.7%	(-2.9, 0.1)	0.13	82	82	1.1 (1.1)	1.4%	(-0.7, 3.0)	0.31	0.08
PY 2	85	85	-0.3 (1.0)	-0.4%	(-2.0, 1.4)	0.76	84	84	0.6 (1.2)	0.7%	(-1.4, 2.7)	0.62	0.57
PY 3	86	88	-2.3** (1.1)	-2.6%	(-4.1, -0.5)	0.04	85	85	1.2 (1.3)	1.4%	(-0.9, 3.3)	0.35	0.04
PY 4	84	88	-4.1*** (1.2)	-4.6%	(-6.1, -2.1)	0.00	81	82	-0.3 (1.3)	-0.4%	(-2.5, 1.8)	0.80	0.03
PY 5	90	91	-1.8 (1.2)	-2.0%	(-3.8, 0.2)	0.14	84	83	2.3* (1.4)	2.8%	(0.0, 4.5)	0.10	0.03
PY 1 through 5	86	87	-1.9** (0.9)	-2.2%	(-3.4, -0.5)	0.03	83	83	1.0 (1.0)	1.2%	(-0.7, 2.7)	0.34	0.03
Acute medical hospitalizations with a complication or comorbidity													
Baseline	49	49	NA	NA	NA	NA	47	49	NA	NA	NA	NA	NA
PY 1	46	46	-0.2 (0.7)	-0.5%	(-1.4, 0.9)	0.76	44	46	-0.6 (0.8)	-1.3%	(-1.9, 0.7)	0.46	0.72
PY 2	45	46	-0.6 (0.7)	-1.2%	(-1.7, 0.6)	0.42	44	46	0.0 (0.8)	-0.1%	(-1.3, 1.3)	0.96	0.62
PY 3	45	45	-0.6 (0.7)	-1.4%	(-1.8, 0.5)	0.36	44	44	1.1 (0.9)	2.6%	(-0.3, 2.6)	0.20	0.11
PY 4	36	38	-1.6** (0.7)	-4.2%	(-2.8, -0.4)	0.03	37	37	0.9 (0.9)	2.4%	(-0.5, 2.3)	0.31	0.03
PY 5	36	37	-1.2 (0.7)	-3.1%	(-2.4, 0.1)	0.12	35	36	0.5 (0.8)	1.5%	(-0.8, 1.8)	0.53	0.13
PY 1 through 5	41	42	-0.8 (0.6)	-1.9%	(-1.8, 0.2)	0.17	41	42	0.4 (0.7)	1.0%	(-0.8, 1.5)	0.58	0.19
Acute medical hospitalizations without any complication or comorbidity													
Baseline	76	77	NA	NA	NA	NA	78	78	NA	NA	NA	NA	NA
PY 1	72	75	-1.1 (0.9)	-1.5%	(-2.6, 0.3)	0.20	73	74	-1.0 (1.0)	-1.3%	(-2.6, 0.6)	0.31	0.92
PY 2	69	71	-1.5 (1.0)	-2.1%	(-3.1, 0.1)	0.13	69	70	-2.1* (1.1)	-3.0%	(-3.9, -0.3)	0.05	0.67
PY 3	68	71	-1.4 (1.0)	-2.1%	(-3.0, 0.2)	0.14	67	69	-2.8** (1.2)	-4.0%	(-4.7, -0.8)	0.02	0.38
PY 4	53	55	-0.4 (1.0)	-0.7%	(-2.0, 1.3)	0.72	52	54	-2.2* (1.3)	-4.0%	(-4.3, -0.1)	0.09	0.27
PY 5	54	55	0.2 (1.0)	0.3%	(-1.5, 1.9)	0.87	51	54	-3.0** (1.2)	-5.5%	(-4.9, -1.0)	0.01	0.05
PY 1 through 5	63	65	-0.9 (0.8)	-1.4%	(-2.2, 0.4)	0.27	62	64	-2.2** (1.0)	-3.5%	(-3.8, -0.6)	0.02	0.30

Table 5.A.1.7b. ***(continued)***

	Track 1 — SSP						Track 1 — Non-SSP						
	CPC+ mean[a]	C mean[a]	Impact estimate[b] (SE)	Percentage impact[c]	90% confidence interval	*p*-Value	CPC+ mean[a]	C mean[a]	Impact estimate[b] (SE)	Percentage impact[c]	90% confidence interval	*p*-Value	*p*-Value for SSP vs. non-SSP difference
Unweighted sample sizes for measures per 1,000 beneficiaries per year[d]													
Number of practices	738	2,979					635	2,264					
Number of beneficiaries	798,817	3,129,830					753,337	2,233,041					
Number of beneficiary-years	3,017,546	11,762,356					2,898,848	8,387,734					

Source: Mathematica's analysis of Medicare claims data from January 2013 through December 2021.

Notes: This table indicates which estimates are statistically significant; when we interpret evidence, we combine evidence from the magnitude of the effect, the *p*-values, findings on related outcomes, subgroups, sensitivity tests, and other data sources about model implementation.

[a] We report the actual, unadjusted averages in the baseline period which are similar for the CPC+ and comparison groups due to matching. In the intervention periods, the comparison group mean is computed by subtracting the regression adjusted difference between the CPC+ and comparison means in each time period from the CPC+ mean in that same time period.

[b] Each impact estimate is regression-adjusted using a difference-in-differences analysis that reflects the difference of the average outcome for Medicare FFS beneficiaries attributed to CPC+ practices in the five years of CPC+ to the average outcome in the baseline year, relative to the same difference over time for Medicare FFS beneficiaries attributed to comparison practices, while controlling for beneficiary characteristics and practice fixed effects.

[c] We calculated percentage impacts relative to what the CPC+ mean would have been in Program Years 1 through 5 (separately and combined) in the absence of the intervention—that is, the unadjusted CPC+ mean minus the impact estimate.

[d] After accounting for weights that adjust for matching and time observed in Medicare FFS, the effective sample sizes fall but are still substantial. For the comparison group, the effective sample size is 43 to 50 percent of the size of the actual comparison group. The effective sample size for the CPC+ group is 96 percent of the actual sample size because it is affected only by time observed (and not by the matching weights).

*/**/*** Significantly different from zero at the 0.10/0.05/0.01 level, two-tailed test.

C = comparison; FFS = fee-for-service; NA = not applicable; PY = Program Year; SE = standard error; SSP = Medicare Shared Savings Program.

Table 5.A.1.8a. Regression-adjusted means and estimated impacts of CPC+ on types of acute hospitalizations for attributed Medicare FFS beneficiaries by program year and average across the five program years, Track 2

	Track 2 — Overall					
	CPC+ mean[a]	C mean[a]	Impact estimate[b] (SE)	Percentage impact[c]	90% confidence interval	*p*-Value
Hospitalizations per 1,000 beneficiaries per year (short-stay acute care and critical access hospitals)						
Acute surgical hospitalizations						
Baseline	90	88	NA	NA	NA	NA
PY 1	88	85	0.8 (0.7)	0.9%	(-0.4, 1.9)	0.27
PY 2	87	84	0.8 (0.7)	1.0%	(-0.3, 1.9)	0.21
PY 3	87	85	0.0 (0.7)	0.0%	(-1.1, 1.1)	0.95
PY 4	71	70	-0.9 (0.7)	-1.2%	(-2.0, 0.2)	0.19
PY 5	68	67	-0.6 (0.7)	-0.8%	(-1.7, 0.6)	0.42
PY 1 through 5	80	78	0.0 (0.6)	0.0%	(-0.9, 1.0)	0.96
Acute surgical hospitalizations with a major complication or comorbidity						
Baseline	22	22	NA	NA	NA	NA
PY 1	22	21	0.1 (0.3)	0.3%	(-0.5, 0.6)	0.83
PY 2	22	22	0.1 (0.3)	0.4%	(-0.4, 0.6)	0.77
PY 3	23	22	-0.2 (0.3)	-0.8%	(-0.7, 0.4)	0.58
PY 4	21	21	-0.5 (0.3)	-2.4%	(-1.0, 0.0)	0.12
PY 5	21	21	-0.5 (0.3)	-2.5%	(-1.1, 0.0)	0.12
PY 1 through 5	22	21	-0.2 (0.3)	-1.0%	(-0.7, 0.2)	0.43
Acute surgical hospitalizations with a complication or comorbidity						
Baseline	22	22	NA	NA	NA	NA
PY 1	21	21	0.2 (0.3)	1.2%	(-0.3, 0.8)	0.46
PY 2	22	21	0.3 (0.3)	1.4%	(-0.2, 0.8)	0.35
PY 3	22	22	0.1 (0.3)	0.6%	(-0.4, 0.7)	0.70
PY 4	20	20	0.3 (0.3)	1.3%	(-0.3, 0.8)	0.43
PY 5	20	19	0.8** (0.3)	3.9%	(0.2, 1.3)	0.02
PY 1 through 5	21	21	0.3 (0.3)	1.6%	(-0.1, 0.8)	0.21
Acute surgical hospitalizations without any complication or comorbidity						
Baseline	46	44	NA	NA	NA	NA
PY 1	45	43	0.4 (0.4)	1.0%	(-0.3, 1.2)	0.33
PY 2	43	41	0.5 (0.4)	1.1%	(-0.3, 1.2)	0.31
PY 3	43	41	0.0 (0.4)	0.0%	(-0.7, 0.7)	0.97
PY 4	30	29	-0.6 (0.4)	-2.1%	(-1.4, 0.1)	0.14
PY 5	27	26	-0.8* (0.4)	-2.8%	(-1.5, -0.1)	0.07

Table 5.A.1.8a. ***(continued)***

	Track 2 — Overall					
	CPC+ mean[a]	C mean[a]	Impact estimate[b] (SE)	Percentage impact[c]	90% confidence interval	*p*-Value
PY 1 through 5	37	36	-0.1 (0.4)	-0.3%	(-0.7, 0.5)	0.78
Acute medical hospitalizations						
Baseline	202	201	NA	NA	NA	NA
PY 1	204	203	-1.3 (1.3)	-0.6%	(-3.5, 0.9)	0.34
PY 2	202	203	-2.4 (1.4)	-1.2%	(-4.7, 0.0)	0.10
PY 3	199	202	-4.8*** (1.5)	-2.3%	(-7.3, -2.3)	0.00
PY 4	174	176	-3.9** (1.6)	-2.2%	(-6.6, -1.3)	0.01
PY 5	178	177	-1.3 (1.6)	-0.7%	(-4.0, 1.4)	0.43
PY 1 through 5	191	191	-2.8** (1.3)	-1.4%	(-4.9, -0.6)	0.03
Acute medical hospitalizations with a major complication or comorbidity						
Baseline	75	74	NA	NA	NA	NA
PY 1	83	82	0.6 (0.8)	0.8%	(-0.6, 1.9)	0.41
PY 2	86	85	-0.2 (0.8)	-0.3%	(-1.6, 1.1)	0.79
PY 3	86	87	-1.4 (0.9)	-1.6%	(-2.9, 0.1)	0.11
PY 4	83	84	-1.9** (1.0)	-2.3%	(-3.6, -0.3)	0.05
PY 5	87	86	0.7 (1.0)	0.8%	(-1.0, 2.4)	0.50
PY 1 through 5	85	85	-0.5 (0.8)	-0.5%	(-1.7, 0.8)	0.53
Acute medical hospitalizations with a complication or comorbidity						
Baseline	49	49	NA	NA	NA	NA
PY 1	46	46	-0.8 (0.5)	-1.7%	(-1.7, 0.1)	0.13
PY 2	45	46	-0.7 (0.6)	-1.6%	(-1.6, 0.2)	0.19
PY 3	44	45	-1.1* (0.6)	-2.3%	(-2.0, -0.1)	0.06
PY 4	37	37	-0.5 (0.6)	-1.4%	(-1.5, 0.4)	0.34
PY 5	36	37	-0.6 (0.6)	-1.7%	(-1.6, 0.3)	0.27
PY 1 through 5	42	42	-0.8 (0.5)	-1.8%	(-1.5, 0.0)	0.10
Acute medical hospitalizations without any complication or comorbidity						
Baseline	78	78	NA	NA	NA	NA
PY 1	75	75	-1.1 (0.7)	-1.5%	(-2.2, 0.0)	0.11
PY 2	71	72	-1.4** (0.7)	-2.0%	(-2.6, -0.2)	0.05
PY 3	69	70	-2.3*** (0.7)	-3.2%	(-3.5, -1.1)	0.00
PY 4	54	55	-1.4* (0.8)	-2.6%	(-2.7, -0.2)	0.06
PY 5	54	55	-1.4* (0.8)	-2.5%	(-2.7, -0.1)	0.08
PY 1 through 5	64	65	-1.5** (0.6)	-2.4%	(-2.6, -0.5)	0.01

Table 5.A.1.8a. ***(continued)***

	Track 2 — Overall					
	CPC+ mean[a]	C mean[a]	Impact estimate[b] (SE)	Percentage impact[c]	90% confidence interval	*p*-Value
Unweighted sample sizes for measures per 1,000 beneficiaries per year[d]						
Number of practices	1,515	3,783				
Number of beneficiaries	1,896,880	4,507,499				
Number of beneficiary-years	7,225,289	17,054,519				

Source: Mathematica's analysis of Medicare claims data from January 2013 through December 2021.

Notes: This table indicates which estimates are statistically significant; when we interpret evidence, we combine evidence from the magnitude of the effect, the *p*-values, findings on related outcomes, subgroups, sensitivity tests, and other data sources about model implementation.

[a] We report the actual, unadjusted averages in the baseline period which are similar for the CPC+ and comparison groups due to matching. In the intervention periods, the comparison group mean is computed by subtracting the regression adjusted difference between the CPC+ and comparison means in each time period from the CPC+ mean in that same time period.

[b] Each impact estimate is regression-adjusted using a difference-in-differences analysis that reflects the difference of the average outcome for Medicare FFS beneficiaries attributed to CPC+ practices in the five years of CPC+ to the average outcome in the baseline year, relative to the same difference over time for Medicare FFS beneficiaries attributed to comparison practices, while controlling for beneficiary characteristics and practice fixed effects.

[c] We calculated percentage impacts relative to what the CPC+ mean would have been in Program Years 1 through 5 (separately and combined) in the absence of the intervention—that is, the unadjusted CPC+ mean minus the impact estimate.

[d] After accounting for weights that adjust for matching and time observed in Medicare FFS, the effective sample sizes fall but are still substantial. For the comparison group, the effective sample size is 40 percent of the size of the actual comparison group. The effective sample size for the CPC+ group is 96 percent of the actual sample size because it is affected only by time observed (and not by the matching weights).

*/**/*** Significantly different from zero at the 0.10/0.05/0.01 level, two-tailed test.

C = comparison; FFS = fee-for-service; NA = not applicable; PY = Program Year; SE = standard error.

Table 5.A.1.8b. Regression-adjusted means and estimated impacts of CPC+ on types of acute hospitalizations for attributed Medicare FFS beneficiaries by program year and average across the five program years, Track 2 by SSP status

	Track 2 — SSP						Track 2 — Non-SSP						
	CPC+ mean[a]	C mean[a]	Impact estimate[b] (SE)	Percentage impact[c]	90% confidence interval	*p*-Value	CPC+ mean[a]	C mean[a]	Impact estimate[b] (SE)	Percentage impact[c]	90% confidence interval	*p*-Value	*p*-Value for SSP vs. non-SSP difference
Hospitalizations per 1,000 beneficiaries per year (short-stay acute care and critical access hospitals)													
Acute surgical hospitalizations													
Baseline	91	88	NA	NA	NA	NA	89	87	NA	NA	NA	NA	NA
PY 1	89	87	-0.5 (1.0)	-0.6%	(-2.2, 1.1)	0.60	88	84	1.8* (0.9)	2.0%	(0.2, 3.3)	0.06	0.09
PY 2	87	85	-0.3 (1.0)	-0.4%	(-2.0, 1.3)	0.75	87	83	1.7* (0.9)	2.1%	(0.3, 3.2)	0.05	0.12
PY 3	88	86	0.0 (0.9)	0.0%	(-1.5, 1.6)	0.99	86	84	-0.1 (0.9)	-0.1%	(-1.7, 1.4)	0.90	0.92
PY 4	71	71	-1.8* (1.0)	-2.4%	(-3.4, -0.1)	0.08	71	69	-0.2 (0.9)	-0.3%	(-1.8, 1.3)	0.79	0.27
PY 5	69	68	-1.6 (1.0)	-2.2%	(-3.3, 0.2)	0.14	68	65	0.4 (1.0)	0.6%	(-1.1, 2.0)	0.66	0.16
PY 1 through 5	80	79	-0.8 (0.8)	-1.0%	(-2.1, 0.5)	0.32	79	77	0.7 (0.8)	0.9%	(-0.6, 2.0)	0.36	0.18
Acute surgical hospitalizations with a major complication or comorbidity													
Baseline	23	22	NA	NA	NA	NA	22	21	NA	NA	NA	NA	NA
PY 1	22	22	-1.2** (0.5)	-5.0%	(-1.9, -0.4)	0.02	22	21	1.0** (0.4)	4.9%	(0.4, 1.7)	0.01	0.00
PY 2	22	22	-1.1** (0.5)	-4.6%	(-1.9, -0.3)	0.03	22	21	1.0** (0.4)	4.7%	(0.3, 1.7)	0.02	0.00
PY 3	23	23	-1.0** (0.5)	-4.4%	(-1.9, -0.2)	0.04	22	22	0.5 (0.5)	2.1%	(-0.3, 1.2)	0.31	0.03
PY 4	21	21	-1.3*** (0.5)	-5.9%	(-2.1, -0.5)	0.01	21	20	0.2 (0.4)	1.2%	(-0.5, 0.9)	0.58	0.02
PY 5	21	22	-1.8*** (0.5)	-7.7%	(-2.7, -0.9)	0.00	21	20	0.6 (0.4)	3.1%	(-0.1, 1.4)	0.16	0.00
PY 1 through 5	22	22	-1.3*** (0.4)	-5.5%	(-1.9, -0.6)	0.00	22	21	0.7* (0.4)	3.2%	(0.1, 1.3)	0.05	0.00
Acute surgical hospitalizations with a complication or comorbidity													
Baseline	22	22	NA	NA	NA	NA	22	22	NA	NA	NA	NA	NA
PY 1	21	21	0.4 (0.5)	2.1%	(-0.3, 1.2)	0.34	21	21	0.1 (0.5)	0.4%	(-0.7, 0.9)	0.84	0.59
PY 2	22	22	0.3 (0.5)	1.6%	(-0.4, 1.1)	0.45	21	21	0.3 (0.4)	1.3%	(-0.4, 1.0)	0.54	0.90
PY 3	23	22	0.5 (0.5)	2.2%	(-0.4, 1.3)	0.34	22	22	-0.2 (0.4)	-0.7%	(-0.9, 0.6)	0.72	0.34

Table 5.A.1.8b. ***(continued)***

	Track 2 — SSP						Track 2 — Non-SSP						
	CPC+ mean[a]	C mean[a]	Impact estimate[b] (SE)	Percentage impact[c]	90% confidence interval	*p*-Value	CPC+ mean[a]	C mean[a]	Impact estimate[b] (SE)	Percentage impact[c]	90% confidence interval	*p*-Value	*p*-Value for SSP vs. non-SSP difference
PY 4	20	20	0.8* (0.5)	3.9%	(0.0, 1.5)	0.10	19	19	-0.1 (0.4)	-0.7%	(-0.9, 0.6)	0.76	0.16
PY 5	21	19	1.3*** (0.5)	6.7%	(0.5, 2.1)	0.01	20	19	0.4 (0.4)	2.1%	(-0.3, 1.1)	0.34	0.18
PY 1 through 5	21	21	0.7* (0.4)	3.2%	(0.0, 1.3)	0.09	21	20	0.1 (0.4)	0.5%	(-0.5, 0.7)	0.80	0.29
Acute surgical hospitalizations without any complication or comorbidity													
Baseline	46	44	NA	NA	NA	NA	46	44	NA	NA	NA	NA	NA
PY 1	46	44	0.2 (0.6)	0.4%	(-0.8, 1.2)	0.77	45	43	0.6 (0.6)	1.4%	(-0.4, 1.7)	0.31	0.61
PY 2	43	41	0.4 (0.6)	1.0%	(-0.7, 1.5)	0.53	43	41	0.5 (0.6)	1.1%	(-0.5, 1.5)	0.43	0.93
PY 3	43	41	0.6 (0.7)	1.4%	(-0.5, 1.6)	0.37	42	41	-0.4 (0.6)	-1.0%	(-1.4, 0.5)	0.47	0.25
PY 4	30	30	-1.2* (0.7)	-4.0%	(-2.3, -0.2)	0.06	31	30	-0.3 (0.6)	-1.1%	(-1.3, 0.6)	0.55	0.32
PY 5	27	26	-1.1* (0.6)	-3.8%	(-2.1, 0.0)	0.10	27	26	-0.6 (0.6)	-2.3%	(-1.6, 0.4)	0.31	0.63
PY 1 through 5	37	36	-0.2 (0.5)	-0.5%	(-1.0, 0.6)	0.70	37	36	0.0 (0.5)	-0.1%	(-0.9, 0.8)	0.92	0.85
Acute medical hospitalizations													
Baseline	209	203	NA	NA	NA	NA	198	199	NA	NA	NA	NA	NA
PY 1	213	207	0.1 (2.0)	0.0%	(-3.3, 3.4)	0.97	197	200	-2.3 (1.8)	-1.2%	(-5.3, 0.6)	0.19	0.37
PY 2	210	204	0.3 (2.2)	0.1%	(-3.3, 3.8)	0.90	196	202	-4.4** (1.9)	-2.2%	(-7.6, -1.3)	0.02	0.10
PY 3	208	204	-2.1 (2.4)	-1.0%	(-6.0, 1.8)	0.38	192	200	-6.9*** (2.0)	-3.5%	(-10.2, -3.7)	0.00	0.12
PY 4	182	178	-2.2 (2.6)	-1.2%	(-6.5, 2.1)	0.41	168	174	-4.8** (1.9)	-2.8%	(-8.0, -1.7)	0.01	0.40
PY 5	188	182	-0.8 (2.7)	-0.4%	(-5.1, 3.6)	0.78	169	173	-2.3 (2.0)	-1.4%	(-5.7, 1.0)	0.25	0.66
PY 1 through 5	199	194	-0.9 (2.1)	-0.5%	(-4.3, 2.4)	0.65	184	189	-4.3*** (1.6)	-2.3%	(-6.9, -1.6)	0.01	0.21
Acute medical hospitalizations with a major complication or comorbidity													
Baseline	76	75	NA	NA	NA	NA	74	74	NA	NA	NA	NA	NA
PY 1	85	84	0.5 (1.1)	0.6%	(-1.3, 2.4)	0.65	82	81	0.7 (1.0)	0.9%	(-1.0, 2.4)	0.50	0.90
PY 2	87	87	-0.5 (1.2)	-0.6%	(-2.5, 1.5)	0.67	85	84	0.0 (1.1)	0.0%	(-1.8, 1.8)	1.00	0.75
PY 3	88	88	-1.7 (1.4)	-1.9%	(-4.0, 0.7)	0.24	84	85	-1.2 (1.1)	-1.4%	(-3.1, 0.6)	0.28	0.81

Table 5.A.1.8b. ***(continued)***

	Track 2 — SSP						Track 2 — Non-SSP						
	CPC+ mean[a]	C mean[a]	Impact estimate[b] (SE)	Percentage impact[c]	90% confidence interval	*p*-Value	CPC+ mean[a]	C mean[a]	Impact estimate[b] (SE)	Percentage impact[c]	90% confidence interval	*p*-Value	*p*-Value for SSP vs. non-SSP difference
PY 4	85	86	-2.1 (1.5)	-2.5%	(-4.6, 0.3)	0.16	81	82	-1.4 (1.2)	-1.7%	(-3.4, 0.6)	0.24	0.72
PY 5	90	90	-0.8 (1.7)	-0.9%	(-3.5, 1.9)	0.62	85	83	1.4 (1.2)	1.7%	(-0.6, 3.5)	0.25	0.27
PY 1 through 5	87	87	-0.9 (1.2)	-1.1%	(-2.9, 1.0)	0.42	83	83	-0.1 (1.0)	-0.2%	(-1.7, 1.4)	0.89	0.59
Acute medical hospitalizations with a complication or comorbidity													
Baseline	51	50	NA	NA	NA	NA	48	48	NA	NA	NA	NA	NA
PY 1	48	47	0.3 (0.8)	0.6%	(-1.0, 1.6)	0.71	44	46	-1.6** (0.7)	-3.6%	(-2.8, -0.5)	0.02	0.07
PY 2	48	46	0.6 (0.8)	1.4%	(-0.7, 2.0)	0.44	44	46	-1.8** (0.7)	-4.0%	(-3.0, -0.6)	0.01	0.03
PY 3	47	45	0.4 (0.8)	0.8%	(-1.0, 1.8)	0.65	42	45	-2.2*** (0.7)	-4.9%	(-3.4, -1.0)	0.00	0.02
PY 4	39	37	0.2 (0.9)	0.4%	(-1.3, 1.7)	0.85	36	37	-1.0 (0.7)	-2.7%	(-2.1, 0.2)	0.15	0.30
PY 5	39	37	0.4 (0.9)	1.2%	(-1.0, 1.9)	0.62	34	36	-1.4* (0.8)	-3.8%	(-2.6, -0.1)	0.07	0.13
PY 1 through 5	44	42	0.4 (0.7)	0.9%	(-0.8, 1.5)	0.58	40	42	-1.6*** (0.6)	-3.9%	(-2.6, -0.6)	0.01	0.03
Acute medical hospitalizations without any complication or comorbidity													
Baseline	82	78	NA	NA	NA	NA	76	77	NA	NA	NA	NA	NA
PY 1	79	76	-0.7 (1.0)	-0.9%	(-2.4, 1.0)	0.49	71	74	-1.4 (0.9)	-1.9%	(-2.9, 0.1)	0.13	0.63
PY 2	75	71	0.1 (1.1)	0.2%	(-1.6, 1.9)	0.89	68	72	-2.6*** (1.0)	-3.7%	(-4.2, -1.0)	0.01	0.06
PY 3	73	70	-0.8 (1.1)	-1.1%	(-2.6, 1.0)	0.47	65	70	-3.5*** (1.0)	-5.1%	(-5.1, -1.8)	0.00	0.07
PY 4	58	54	-0.2 (1.2)	-0.3%	(-2.1, 1.8)	0.87	51	55	-2.4** (1.0)	-4.5%	(-4.0, -0.8)	0.02	0.15
PY 5	59	55	-0.4 (1.2)	-0.6%	(-2.4, 1.7)	0.76	51	55	-2.4** (1.0)	-4.5%	(-4.0, -0.7)	0.02	0.22
PY 1 through 5	68	65	-0.4 (1.0)	-0.6%	(-2.0, 1.2)	0.69	60	65	-2.5*** (0.8)	-4.0%	(-3.8, -1.1)	0.00	0.09
Unweighted sample sizes for measures per 1,000 beneficiaries per year[d]													
Number of practices	636	1,817					879	1,966					
Number of beneficiaries	847,208	2,257,322					1,053,634	2,261,852					
Number of beneficiary-years	3,204,963	8,538,135					4,020,326	8,516,384					

Table 5.A.1.8b. ***(continued)***

Source: Mathematica's analysis of Medicare claims data from January 2013 through December 2021.

Notes: This table indicates which estimates are statistically significant; when we interpret evidence, we combine evidence from the magnitude of the effect, the *p*-values, findings on related outcomes, subgroups, sensitivity tests, and other data sources about model implementation.

[a] We report the actual, unadjusted averages in the baseline period which are similar for the CPC+ and comparison groups due to matching. In the intervention periods, the comparison group mean is computed by subtracting the regression adjusted difference between the CPC+ and comparison means in each time period from the CPC+ mean in that same time period.

[b] Each impact estimate is regression-adjusted using a difference-in-differences analysis that reflects the difference of the average outcome for Medicare FFS beneficiaries attributed to CPC+ practices in the five years of CPC+ to the average outcome in the baseline year, relative to the same difference over time for Medicare FFS beneficiaries attributed to comparison practices, while controlling for beneficiary characteristics and practice fixed effects.

[c] We calculated percentage impacts relative to what the CPC+ mean would have been in Program Years 1 through 5 (separately and combined) in the absence of the intervention—that is, the unadjusted CPC+ mean minus the impact estimate.

[d] After accounting for weights that adjust for matching and time observed in Medicare FFS, the effective sample sizes fall but are still substantial. For the comparison group, the effective sample size is 38 to 43 percent of the size of the actual comparison group. The effective sample size for the CPC+ group is 96 percent of the actual sample size because it is affected only by time observed (and not by the matching weights).

*/**/*** Significantly different from zero at the 0.10/0.05/0.01 level, two-tailed test.

C = comparison; FFS = fee-for-service; NA = not applicable; PY = Program Year; SE = standard error; SSP = Medicare Shared Savings Program.

Table 5.A.1.9a. Estimated average annual impacts of CPC+ on acute hospitalizations across the five program years, by baseline practice characteristics, Track 1

Practice subgroup definition, based on baseline characteristics	Track 1 – Overall			
	Number (percentage) of CPC+ beneficiaries in subgroup at baseline	Impact estimate (standard error)	Percentage impact	*p*-Value for difference in impact estimates between subgroups[a]
Main analysis (all practices)- average annual estimate for PY 1 through PY 5	**-**	**-2.5* (1.4)**	**-0.9%**	**-**
Whether practice participated in prior primary care transformation initiatives (recognized as a medical home or participated in MAPCP or CPC Classic)				
Yes	468,487 (53.6%)	-2.4 (2.0)	-0.9%	
No	405,383 (46.4%)	-2.9 (1.9)	-1.0%	0.95
Large and medium versus small practice based on number of primary care practitioners				
Large (6+ primary care practitioners)	404,456 (46.3%)	-5.2** (2.1)	-1.9%	
Medium (3–5 primary care practitioners)	282,380 (32.3%)	2.1 (2.7)	0.8%	
Small (1–2 primary care practitioners)	187,034 (21.4%)	-4.2 (2.9)	-1.5%	0.08
Whether hospital- or system-owned versus independent (based on IQVIA data)[b]				
Hospital- or system-owned	474,606 (54.3%)	-1.5 (2.0)	-0.5%	
Independent	399,264 (45.7%)	-4.0 (2.0)	-1.5%	0.43
Whether the practice shared a TIN with another primary care practice[b]				
Shared a TIN with another primary care practice	684,507 (78.3%)	-2.3 (1.6)	-0.8%	
Did not share a TIN with another primary care practice	189,364 (21.7%)	-3.2 (2.7)	-1.2%	0.79
Practice type: multi-specialty versus primary care only				
Multi-specialty	170,691 (19.5%)	1.0 (3.9)	0.4%	
Primary care only	703,179 (80.5%)	-3.5 (1.5)	-1.3%	0.21
Urbanicity of practice's county: rural or suburban location versus urban location				
Rural	89,834 (10.3%)	4.8 (5.0)	2.0%	
Suburban	156,799 (17.9%)	-2.1 (4.2)	-0.8%	
Urban	627,237 (71.8%)	-3.8 (1.5)	-1.4%	0.31

Source: Mathematica's analysis of Medicare claims data from January 2013 through December 2021.

Note: The estimates (and standard errors) in the impact estimate column show subgroup-specific impacts over the five years of CPC+, separately, for each practice characteristic listed in the table. We only tested differences within each subgroup if the estimates were significantly different between the two subgroups (that is, the *p*-value in the last column was <.10). Asterisks denote whether the impact estimate within a subgroup was significantly different from zero when estimates were significantly different between the subgroup categories.

[a] The *p*-values in the last column represent results from testing for statistically significant differences in impact estimates between the subgroups, based on the baseline practice characteristic (using a t-test for subgroups with two categories and from an F-test for subgroups with more than two categories).

[b] Since ownership status of a practice is likely to be highly correlated with whether the practice shares TIN with other practices, we included only one of these characteristics at a time in these regressions.

*/**/***Within-subgroup estimate significantly different from zero at the 0.10/0.05/0.01 level, two-tailed test.

CPC = Comprehensive Primary Care; MAPCP = Multi-payer Advanced Primary Care Practice Demonstration; PY = Program Year; TIN = Tax Identification Number

Table 5.A.1.9b. Estimated average annual impacts of CPC+ on acute hospitalizations across the five program years, by baseline practice characteristics and SSP status, Track 1

Practice subgroup definition, based on baseline characteristics	Track 1 – SSP				Track 1 – Non-SSP			
	Number (percentage) of CPC+ beneficiaries in subgroup at baseline	Impact estimate (standard error)	Percentage impact	*p*-Value for difference in impact estimates between subgroups[a]	Number (percentage) of CPC+ beneficiaries in subgroup at baseline	Impact estimate (standard error)	Percentage impact	*p*-Value for difference in impact estimates between subgroups[a]
Main analysis (all practices)- average annual estimate for PY 1 through PY 5	**-**	**-4.5** (1.8)**	**-1.6%**	**-**	**-**	**-0.1 (2.2)**	**0.0%**	**-**
Whether practice participated in prior primary care transformation initiatives (recognized as a medical home or participated in MAPCP or CPC Classic)								
Yes	214,075 (47.7%)	-1.1 (2.6)	-0.4%		254,262 (59.8%)	-3.6 (2.9)	-1.4%	
No	234,948 (52.3%)	-7.6*** (2.5)	-2.7%	0.05	170,586 (40.2%)	3.7 (3.0)	1.4%	0.10
Large and medium versus small practice based on number of primary care practitioners								
Large (6+ primary care practitioners)	189,229 (42.1%)	-6.4 (2.9)	-2.4%		215,122 (50.6%)	-4.5 (3.0)	-1.7%	
Medium (3–5 primary care practitioners)	156,338 (34.8%)	-0.8 (2.9)	-0.3%		126,106 (29.7%)	5.9 (4.4)	2.3%	
Small (1–2 primary care practitioners)	103,455 (23.0%)	-6.7 (3.7)	-2.4%	0.49	83,621 (19.7%)	-0.7 (4.4)	-0.3%	0.19
Whether hospital- or system-owned versus independent (based on IQVIA data)[b]								
Hospital- or system-owned	250,558 (55.8%)	-4.3 (2.4)	-1.6%		224,086 (52.7%)	1.6 (3.2)	0.6%	
Independent	198,464 (44.2%)	-4.8 (2.8)	-1.7%	0.65	200,762 (47.3%)	-3.2 (2.8)	-1.3%	0.15
Whether the practice shared a TIN with another primary care practice[b]								
Shared a TIN with another primary care practice	366,843 (81.7%)	-4.1 (2.0)	-1.5%		317,749 (74.8%)	0.1 (2.6)	0.0%	
Did not share a TIN with another primary care practice	82,179 (18.3%)	-6.3 (4.4)	-2.3%	0.77	107,099 (25.2%)	-1.5 (3.5)	-0.6%	0.56
Practice type: multi-specialty versus primary care only								
Multi-specialty	76,547 (17.0%)	-8.6 (4.1)	-3.1%		94,082 (22.1%)	8.9 (6.1)	3.7%	
Primary care only	372,475 (83.0%)	-3.7 (2.0)	-1.3%	0.30	330,766 (77.9%)	-3.4 (2.2)	-1.2%	0.04
Urbanicity of practice's county: rural or suburban location versus urban location								
Rural	22,327 (5.0%)	-3.8 (10.2)	-1.5%		67,372 (15.9%)	8.2 (5.7)	3.4%	
Suburban	74,982 (16.7%)	-2.2 (4.7)	-0.8%		81,785 (19.3%)	-2.6 (6.3)	-1.0%	
Urban	351,712 (78.3%)	-5.0 (2.0)	-1.8%	0.82	275,691 (64.9%)	-2.3 (2.5)	-0.8%	0.53

Source: Mathematica's analysis of Medicare claims data from January 2013 through December 2021.

Note: The estimates (and standard errors) in the impact estimate column show subgrou*p*-specific impacts over the five years of CPC+, separately, for each practice characteristic listed in the table. We only tested differences within each subgroup if the estimates were significantly different between the two subgroups (that is, the *p*-value in the last

Table 5.A.1.9b. ***(continued)***

column was <.10). Asterisks denote whether the impact estimate within a subgroup was significantly different from zero when estimates were significantly different between the subgroup categories.

[a] The *p*-values in the last column represent results from testing for statistically significant differences in impact estimates between the subgroups, based on the baseline practice characteristic (using a t-test for subgroups with two categories and from an F-test for subgroups with more than two categories).

[b] Since ownership status of a practice is likely to be highly correlated with whether the practice shares TIN with other practices, we included only one of these characteristics at a time in these regressions.

*/**/***Within-subgroup estimate significantly different from zero at the 0.10/0.05/0.01 level, two-tailed test.

CPC = Comprehensive Primary Care; MAPCP = Multi-payer Advanced Primary Care Practice Demonstration; PY = Program Year; SSP = Medicare Shared Savings Program; TIN = Tax Identification Number

Table 5.A.1.10a. Estimated average annual impacts of CPC+ on acute hospitalizations across the five program years, by baseline practice characteristics, Track 2

Practice subgroup definition, based on baseline characteristics	Track 2 – Overall			
	Number (percentage) of CPC+ beneficiaries in subgroup at baseline	Impact estimate (standard error)	Percentage impact	*p*-Value for difference in impact estimates between subgroups[a]
Main analysis (all practices)- average annual estimate for PY 1 through PY 5	**-**	**-2.7* (1.6)**	**-1.0%**	**-**
Whether practice participated in prior primary care transformation initiatives (recognized as a medical home or participated in MAPCP or CPC Classic)				
Yes	865,798 (81.2%)	-4.1 (1.8)	-1.5%	
No	201,028 (18.8%)	1.7 (2.8)	0.6%	0.10
Large and medium versus small practice based on number of primary care practitioners				
Large (6+ primary care practitioners)	589,224 (55.2%)	-5.8 (2.1)	-2.1%	
Medium (3–5 primary care practitioners)	340,406 (31.9%)	0.7 (2.7)	0.3%	
Small (1–2 primary care practitioners)	137,196 (12.9%)	-0.3 (3.7)	-0.1%	0.26
Whether hospital- or system-owned versus independent (based on IQVIA data)[b]				
Hospital- or system-owned	619,957 (58.1%)	0.1 (2.0)	0.0%	
Independent	446,869 (41.9%)	-7.3*** (2.3)	-2.7%	0.01
Whether the practice shared a TIN with another primary care practice[b]				
Shared a TIN with another primary care practice	913,196 (85.6%)	-2.2 (1.7)	-0.8%	
Did not share a TIN with another primary care practice	153,630 (14.4%)	-6.0 (3.6)	-2.3%	0.16
Practice type: multi-specialty versus primary care only				
Multi-specialty	278,801 (26.1%)	-4.4 (3.6)	-1.6%	
Primary care only	788,025 (73.9%)	-2.5 (1.7)	-0.9%	0.86
Urbanicity of practice's county: rural or suburban location versus urban location				
Rural	82,613 (7.7%)	-3.3 (4.7)	-1.3%	
Suburban	170,323 (16.0%)	-5.6 (4.7)	-2.1%	
Urban	813,890 (76.3%)	-2.4 (1.7)	-0.9%	0.80

Source: Mathematica's analysis of Medicare claims data from January 2013 through December 2021.

Note: The estimates (and standard errors) in the impact estimate column show subgroup*p*-specific impacts over the five years of CPC+, separately, for each practice characteristic listed in the table. We only tested differences within each subgroup if the estimates were significantly different between the two subgroups (that is, the *p*-value in the last column was <.10). Asterisks denote whether the impact estimate within a subgroup was significantly different from zero when estimates were significantly different between the subgroup categories.

[a] The *p*-values in the last column represent results from testing for statistically significant differences in impact estimates between the subgroups, based on the baseline practice characteristic (using a t-test for subgroups with two categories and from an F-test for subgroups with more than two categories).

[b] Since ownership status of a practice is likely to be highly correlated with whether the practice shares TIN with other practices, we included only one of these characteristics at a time in these regressions.

*/**/***Within-subgroup estimate significantly different from zero at the 0.10/0.05/0.01 level, two-tailed test.

CPC = Comprehensive Primary Care; MAPCP = Multi-payer Advanced Primary Care Practice Demonstration; PY = Program Year; TIN = Tax Identification Number

Table 5.A.1.10b. Estimated average annual impacts of CPC+ on acute hospitalizations across the five program years, by baseline practice characteristics and SSP status, Track 2

Practice subgroup definition, based on baseline characteristics	Track 2 – SSP				Track 2 – Non-SSP			
	Number (percentage) of CPC+ beneficiaries in subgroup at baseline	Impact estimate (standard error)	Percentage impact	*p*-Value for difference in impact estimates between subgroups[a]	Number (percentage) of CPC+ beneficiaries in subgroup at baseline	Impact estimate (standard error)	Percentage impact	*p*-Value for difference in impact estimates between subgroups[a]
Main analysis (all practices)- average annual estimate for PY 1 through PY 5	**-**	**-1.7 (2.4)**	**-0.6%**	**-**	**-**	**-3.5* (2.0)**	**-1.3%**	**-**
Whether practice participated in prior primary care transformation initiatives (recognized as a medical home or participated in MAPCP or CPC Classic)								
Yes	385,875 (81.8%)	-4.8* (2.6)	-1.7%		479,947 (80.6%)	-2.8 (2.4)	-1.0%	
No	85,762 (18.2%)	12.2** (4.8)	4.2%	0.00	115,242 (19.4%)	-6.5 (3.3)	-2.5%	0.34
Large and medium versus small practice based on number of primary care practitioners								
Large (6+ primary care practitioners)	279,067 (59.2%)	-4.2 (3.2)	-1.4%		310,301 (52.1%)	-6.0 (2.7)	-2.3%	
Medium (3–5 primary care practitioners)	134,103 (28.4%)	0.7 (3.8)	0.3%		206,177 (34.6%)	0.5 (3.6)	0.2%	
Small (1–2 primary care practitioners)	58,467 (12.4%)	4.2 (5.6)	1.5%	1.00	78,712 (13.2%)	-4.1 (4.8)	-1.5%	0.14
Whether hospital- or system-owned versus independent (based on IQVIA data)[b]								
Hospital- or system-owned	289,350 (61.4%)	0.2 (2.8)	0.1%		330,724 (55.6%)	0.3 (2.8)	0.1%	
Independent	182,287 (38.6%)	-4.9 (3.7)	-1.8%	0.11	264,465 (44.4%)	-8.2*** (2.9)	-3.2%	0.03
Whether the practice shared a TIN with another primary care practice[b]								
Shared a TIN with another primary care practice	416,348 (88.3%)	-0.5 (2.4)	-0.2%		496,945 (83.5%)	-2.9 (2.3)	-1.1%	
Did not share a TIN with another primary care practice	55,289 (11.7%)	-8.8 (6.4)	-3.2%	0.06	98,244 (16.5%)	-5.2 (4.2)	-2.0%	0.79
Practice type: multi-specialty versus primary care only								
Multi-specialty	116,601 (24.7%)	-9.1 (5.3)	-3.1%		162,149 (27.2%)	0.8 (5.1)	0.3%	
Primary care only	355,036 (75.3%)	0.7 (2.5)	0.2%	0.11	433,040 (72.8%)	-5.1 (2.2)	-1.9%	0.13
Urbanicity of practice's county: rural or suburban location versus urban location								
Rural	18,533 (3.9%)	-6.1 (7.7)	-2.4%		63,941 (10.7%)	-2.9 (5.8)	-1.2%	
Suburban	75,938 (16.1%)	-9.2 (6.7)	-3.3%		94,390 (15.9%)	0.5 (6.9)	0.2%	
Urban	377,166 (80.0%)	0.0 (2.5)	0.0%	0.30	436,858 (73.4%)	-4.5 (2.3)	-1.6%	0.80

Source: Mathematica's analysis of Medicare claims data from January 2013 through December 2021.

Note: The estimates (and standard errors) in the impact estimate column show subgrou*p*-specific impacts over the five years of CPC+, separately, for each practice characteristic listed in the table. We only tested differences within each subgroup if the estimates were significantly different between the two subgroups (that is, the *p*-value in the last

Table 5.A.1.10b. ***(continued)***

column was <.10). Asterisks denote whether the impact estimate within a subgroup was significantly different from zero when estimates were significantly different between the subgroup categories.

[a] The *p*-values in the last column represent results from testing for statistically significant differences in impact estimates between the subgroups, based on the baseline practice characteristic (using a t-test for subgroups with two categories and from an F-test for subgroups with more than two categories).

[b] Since ownership status of a practice is likely to be highly correlated with whether the practice shares TIN with other practices, we included only one of these characteristics at a time in these regressions.

*/**/***Within-subgroup estimate significantly different from zero at the 0.10/0.05/0.01 level, two-tailed test.

CPC = Comprehensive Primary Care; MAPCP = Multi-payer Advanced Primary Care Practice Demonstration; PY = Program Year; SSP = Medicare Shared Savings Program; TIN = Tax Identification Number

Table 5.A.1.11a. Estimated average annual impacts of CPC+ on acute hospitalizations across the five program years, by baseline beneficiary characteristics, Track 1

	Track 1 – Overall			
Beneficiary subgroup definition, based on baseline characteristics	**Number (percentage) of CPC+ beneficiaries in subgroup at baseline**	**Impact estimate (standard error)**	**Percentage impact**	***p*-Value for difference in impact estimates between subgroups[a]**
Main analysis (all beneficiaries)- average annual estimate for PY 1 through PY 5	**-**	**-2.5* (1.4)**	**-0.9%**	**-**
Patients in the highest quartile of the HCC score distribution				
Yes	203,811 (25.9%)	-2.9 (4.4)	-0.5%	
No	583,156 (74.1%)	-2.0 (1.1)	-1.1%	0.82
Patients in the highest decile of the HCC score distribution or who have dementia				
Yes	123,085 (15.6%)	-6.5 (6.1)	-0.9%	
No	663,882 (84.4%)	-1.6 (1.2)	-0.7%	0.42
Patients with anxiety/depression or substance use disorders				
Yes	120,562 (16.6%)	-1.3 (4.4)	-0.3%	
No	604,012 (83.4%)	-1.0 (1.4)	-0.4%	0.95
Patients with multiple chronic conditions (at least 2 of 12 frequently occurring chronic conditions[b]) and one or more hospitalizations[c]				
Yes	68,204 (8.7%)	-2.3 (8.9)	-0.3%	
No	718,763 (91.3%)	-2.4 (1.3)	-1.0%	0.99
Patients dually eligible for Medicare and Medicaid				
Yes	107,885 (12.6%)	-4.3 (4.6)	-1.0%	
No	746,776 (87.4%)	-2.2 (1.4)	-0.9%	0.64

Source: Mathematica's analysis of Medicare claims data from January 2013 through December 2021.

Note: Beneficiary characteristics to determine subgroup membership are measured at the start of the year-long baseline period for baseline observations and at the start of Program Year 1 for observations in the intervention period (Program Years 1 through 5). The estimates (and standard errors) in the impact estimate column show subgrou*p*-specific impacts, separately for each beneficiary characteristic listed in the table. We only tested differences *within* each subgroup if the estimates were significantly different *between* the two subgroups (that is, the *p*-value in the last column was <.10). Asterisks denote whether the impact estimate *within* a subgroup was significantly different from zero when estimates were significantly different between the subgroup categories. Because we could not observe diagnoses (which are used to determine HCCs and calculate HCC scores) at baseline for beneficiaries who were new to Medicare during the program years, we excluded new Medicare beneficiaries from all subgroup analyses (except the analysis based on dual status since beneficiaries who are new to Medicare, by definition, could not have been enrolled in both Medicare and Medicaid prior to joining Medicare). Due to this process, about 20 percent of observations from the regressions were excluded for the subgroups defined by HCC score and chronic conditions. Therefore, the main impact estimate of -2.5 for Track 1 overall may not lie between the impact estimates for these subgroups.

[a] The *p*-values in the last column represent results from testing for statistically significant differences in impact estimates between the subgroups, based on the baseline beneficiary characteristic (using a t-test for all subgroups).

[b] The 12 frequently occurring chronic conditions are congestive heart failure, chronic obstructive pulmonary disease, history of acute myocardial infarction, ischemic heart disease, diabetes, metastatic cancer and acute leukemia, history of stroke, depression, dementia, atrial fibrillation, rheumatoid arthritis or osteoarthritis, and chronic kidney disease.

[c] For observations in the baseline year, hospitalizations are measured in 2015, the year before the start of the baseline year. For observations in the intervention period, hospitalizations are measured in 2016, the year before the start of Program Year 1.

*/**/*** Significantly different from zero at the 0.10/0.05/0.01 level, two-tailed test.

HCC = hierarchical condition category; PY = Program Year.

Table 5.A.1.11b. Estimated average annual impacts of CPC+ on acute hospitalizations across the five program years, by baseline beneficiary characteristics and SSP status, Track 1

	Track 1 – SSP				Track 1 – Non-SSP			
Beneficiary subgroup definition, based on baseline characteristics	Number (percentage) of CPC+ beneficiaries in subgroup at baseline	Impact estimate (standard error)	Percentage impact	*p*-Value for difference in impact estimates between subgroups[a]	Number (percentage) of CPC+ beneficiaries in subgroup at baseline	Impact estimate (standard error)	Percentage impact	*p*-Value for difference in impact estimates between subgroups[a]
Main analysis (all beneficiaries)- average annual estimate for PY 1 through PY 5	**-**	**-4.5** (1.8)**	**-1.6%**	**-**	**-**	**-0.1 (2.2)**	**0.0%**	**-**
Patients in the highest quartile of the HCC score distribution								
Yes	115,215 (26.8%)	-4.7 (5.3)	-0.8%		88,864 (25.0%)	-0.5 (7.3)	-0.1%	
No	315,425 (73.2%)	-3.6 (1.5)	-1.9%	0.83	266,666 (75.0%)	0.0 (1.6)	0.0%	0.95
Patients in the highest decile of the HCC score distribution or who have dementia								
Yes	68,759 (16.0%)	-4.1 (7.5)	-0.6%		54,382 (15.3%)	-8.5 (9.7)	-1.1%	
No	361,881 (84.0%)	-4.0 (1.7)	-1.8%	0.99	301,148 (84.7%)	1.1 (1.8)	0.5%	0.32
Patients with anxiety/depression or substance use disorders								
Yes	66,746 (16.8%)	-4.1 (5.7)	-1.0%		53,792 (16.4%)	2.2 (6.7)	0.5%	
No	329,703 (83.2%)	-2.4 (1.8)	-1.0%	0.78	273,568 (83.6%)	0.7 (2.1)	0.3%	0.82
Patients with multiple chronic conditions (at least 2 of 12 frequently occurring chronic conditions[b]) and one or more hospitalizations[c]								
Yes	38,153 (8.9%)	0.3 (11.0)	0.0%		30,089 (8.5%)	-4.5 (14.1)	-0.5%	
No	392,487 (91.1%)	-4.5 (1.7)	-1.8%	0.66	325,442 (91.5%)	0.1 (1.9)	0.0%	0.74
Patients dually eligible for Medicare and Medicaid								
Yes	55,728 (11.9%)	-15.1** (6.3)	-3.5%		51,626 (13.3%)	6.8 (6.6)	1.7%	
No	410,653 (88.1%)	-3.1* (1.8)	-1.2%	0.06	335,619 (86.7%)	-1.0 (2.2)	-0.4%	0.24

Source: Mathematica's analysis of Medicare claims data from January 2013 through December 2021.

Note: Beneficiary characteristics to determine subgroup membership are measured at the start of the year-long baseline period for baseline observations and at the start of Program Year 1 for observations in the intervention period (Program Years 1 through 5). The estimates (and standard errors) in the impact estimate column show subgroup-specific impacts, separately for each beneficiary characteristic listed in the table. We only tested differences *within* each subgroup if the estimates were significantly different *between* the two subgroups (that is, the *p*-value in the last column was <.10). Asterisks denote whether the impact estimate *within* a subgroup was significantly different from zero when estimates were significantly different between the subgroup categories. Because we could not observe diagnoses (which are used to determine HCCs and calculate HCC scores) at baseline for beneficiaries who were new to Medicare during the program years, we excluded new Medicare beneficiaries from all subgroup analyses (except the analysis based on dual status since beneficiaries who are new to Medicare, by definition, could not have been enrolled in both Medicare and Medicaid prior to joining Medicare). Due to this process, about 20 percent of observations from the regressions were excluded for the subgroups defined by HCC score and chronic conditions. Therefore, the main impact estimate of -4.5 for Track 1 SSP and -0.1 for Track 1 Non-SSP may not lie between the impact estimates for these subgroups.

[a] The *p*-values in the last column represent results from testing for statistically significant differences in impact estimates between the subgroups, based on the baseline beneficiary characteristic (using a t-test for all subgroups).

[b] The 12 frequently occurring chronic conditions are congestive heart failure, chronic obstructive pulmonary disease, history of acute myocardial infarction, ischemic heart disease, diabetes, metastatic cancer and acute leukemia, history of stroke, depression, dementia, atrial fibrillation, rheumatoid arthritis or osteoarthritis, and chronic kidney disease.

Table 5.A.1.11b. ***(continued)***

[c] For observations in the baseline year, hospitalizations are measured in 2015, the year before the start of the baseline year. For observations in the intervention period, hospitalizations are measured in 2016, the year before the start of Program Year 1.

*/**/*** Significantly different from zero at the 0.10/0.05/0.01 level, two-tailed test.

HCC = hierarchical condition category; PY = Program Year; SSP = Medicare Shared Savings Program.

Table 5.A.1.12a. Estimated average annual impacts of CPC+ on acute hospitalizations across the five program years, by baseline beneficiary characteristics, Track 2

	Track 2 – Overall			
Beneficiary subgroup definition, based on baseline characteristics	**Number (percentage) of CPC+ beneficiaries in subgroup at baseline**	**Impact estimate (standard error)**	**Percentage impact**	***p*-Value for difference in impact estimates between subgroups[a]**
Main analysis (all beneficiaries)- average annual estimate for PY 1 through PY 5	**-**	**-2.7* (1.6)**	**-1.0%**	**-**
Patients in the highest quartile of the HCC score distribution				
Yes	268,430 (26.1%)	-3.1 (4.5)	-0.5%	
No	761,970 (73.9%)	-0.9 (1.2)	-0.5%	0.62
Patients in the highest decile of the HCC score distribution or who have dementia				
Yes	162,510 (15.8%)	-3.9 (6.3)	-0.5%	
No	867,891 (84.2%)	-1.1 (1.3)	-0.5%	0.65
Patients with anxiety/depression or substance use disorders				
Yes	164,048 (17.3%)	-5.2 (4.5)	-1.3%	
No	784,877 (82.7%)	-0.4 (1.4)	-0.2%	0.29
Patients with multiple chronic conditions (at least 2 of 12 frequently occurring chronic conditions[b]) and one or more hospitalizations[c]				
Yes	90,543 (8.8%)	1.7 (9.4)	0.2%	
No	939,858 (91.2%)	-1.8 (1.4)	-0.7%	0.71
Patients dually eligible for Medicare and Medicaid				
Yes	140,782 (12.5%)	-10.0** (4.9)	-2.3%	
No	984,688 (87.5%)	-1.5 (1.5)	-0.6%	0.08

Source: Mathematica's analysis of Medicare claims data from January 2013 through December 2021.

Note: Beneficiary characteristics to determine subgroup membership are measured at the start of the year-long baseline period for baseline observations and at the start of Program Year 1 for observations in the intervention period (Program Years 1 through 5). The estimates (and standard errors) in the impact estimate column show subgroup-specific impacts, separately for each beneficiary characteristic listed in the table. We only tested differences *within* each subgroup if the estimates were significantly different *between* the two subgroups (that is, the p-value in the last column was <.10). Asterisks denote whether the impact estimate *within* a subgroup was significantly different from zero when estimates were significantly different between the subgroup categories. Because we could not observe diagnoses (which are used to determine HCCs and calculate HCC scores) at baseline for beneficiaries who were new to Medicare during the program years, we excluded new Medicare beneficiaries from all subgroup analyses (except the analysis based on dual status since beneficiaries who are new to Medicare, by definition, could not have been enrolled in both Medicare and Medicaid prior to joining Medicare). Due to this process, about 20 percent of observations from the regressions were excluded for the subgroups defined by HCC score and chronic conditions. Therefore, the main impact estimate of -2.7 for Track 2 overall may not lie between the impact estimates for these subgroups.

[a] The p-values in the last column represent results from testing for statistically significant differences in impact estimates between the subgroups, based on the baseline beneficiary characteristic (using a t-test for all subgroups).

[b] The 12 frequently occurring chronic conditions are congestive heart failure, chronic obstructive pulmonary disease, history of acute myocardial infarction, ischemic heart disease, diabetes, metastatic cancer and acute leukemia, history of stroke, depression, dementia, atrial fibrillation, rheumatoid arthritis or osteoarthritis, and chronic kidney disease.

[c] For observations in the baseline year, hospitalizations are measured in 2015, the year before the start of the baseline year. For observations in the intervention period, hospitalizations are measured in 2016, the year before the start of Program Year 1.

*/**/*** Significantly different from zero at the 0.10/0.05/0.01 level, two-tailed test.

HCC = hierarchical condition category; PY = Program Year.

Table 5.A.1.12b. Estimated average annual impacts of CPC+ on acute hospitalizations across the five program years, by baseline beneficiary characteristics and SSP status, Track 2

	Track 2 – SSP				Track 2 – Non-SSP			
Beneficiary subgroup definition, based on baseline characteristics	Number (percentage) of CPC+ beneficiaries in subgroup at baseline	Impact estimate (standard error)	Percentage impact	*p*-Value for difference in impact estimates between subgroups[a]	Number (percentage) of CPC+ beneficiaries in subgroup at baseline	Impact estimate (standard error)	Percentage impact	*p*-Value for difference in impact estimates between subgroups[a]
Main analysis (all beneficiaries)- average annual estimate for PY 1 through PY 5	**-**	**-1.7 (2.4)**	**-0.6%**	**-**	**-**	**-3.5* (2.0)**	**-1.3%**	**-**
Patients in the highest quartile of the HCC score distribution								
Yes	120,947 (26.8%)	-0.6 (6.6)	-0.1%		146,522 (25.5%)	-5.4 (6.0)	-0.9%	
No	330,277 (73.2%)	1.2 (2.0)	0.6%	0.78	428,947 (74.5%)	-2.5 (1.5)	-1.4%	0.63
Patients in the highest decile of the HCC score distribution or who have dementia								
Yes	71,030 (15.7%)	-6.3 (9.7)	-0.8%		90,895 (15.8%)	-2.3 (8.3)	-0.3%	
No	380,194 (84.3%)	2.0 (2.1)	0.9%	0.39	484,574 (84.2%)	-3.6 (1.6)	-1.6%	0.88
Patients with anxiety/depression or substance use disorders								
Yes	74,382 (17.8%)	-4.1 (6.5)	-1.0%		89,058 (16.8%)	-6.3 (6.2)	-1.5%	
No	342,453 (82.2%)	2.7 (2.2)	1.1%	0.28	439,501 (83.2%)	-2.9 (1.8)	-1.2%	0.59
Patients with multiple chronic conditions (at least 2 of 12 frequently occurring chronic conditions[b]) and one or more hospitalizations[c]								
Yes	41,080 (9.1%)	0.6 (13.8)	0.1%		49,139 (8.5%)	2.5 (12.7)	0.3%	
No	410,144 (90.9%)	0.8 (2.3)	0.3%	0.99	526,331 (91.5%)	-3.8 (1.7)	-1.6%	0.62
Patients dually eligible for Medicare and Medicaid								
Yes	55,837 (11.3%)	-15.2** (7.7)	-3.4%		84,414 (13.5%)	-6.3 (6.2)	-1.5%	
No	438,154 (88.7%)	0.3 (2.4)	0.1%	0.04	542,895 (86.5%)	-2.9 (2.0)	-1.2%	0.59

Source: Mathematica's analysis of Medicare claims data from January 2013 through December 2021.

Note: Beneficiary characteristics to determine subgroup membership are measured at the start of the year-long baseline period for baseline observations and at the start of Program Year 1 for observations in the intervention period (Program Years 1 through 5). The estimates (and standard errors) in the impact estimate column show subgroup-specific impacts, separately for each beneficiary characteristic listed in the table. We only tested differences *within* each subgroup if the estimates were significantly different *between* the two subgroups (that is, the p-value in the last column was <.10). Asterisks denote whether the impact estimate *within* a subgroup was significantly different from zero when estimates were significantly different between the subgroup categories. Because we could not observe diagnoses (which are used to determine HCCs and calculate HCC scores) at baseline for beneficiaries who were new to Medicare during the program years, we excluded new Medicare beneficiaries from all subgroup analyses (except the analysis based on dual status since beneficiaries who are new to Medicare, by definition, could not have been enrolled in both Medicare and Medicaid prior to joining Medicare). Due to this process, about 20 percent of observations from the regressions were excluded for the subgroups defined by HCC score and chronic conditions. Therefore, the main impact estimate of -1.7 for Track 2 SSP and -3.5 for Track 2 Non-SSP may not lie between the impact estimates for these subgroups.

[a] The p-values in the last column represent results from testing for statistically significant differences in impact estimates between the subgroups, based on the baseline beneficiary characteristic (using a t-test for all subgroups).

[b] The 12 frequently occurring chronic conditions are congestive heart failure, chronic obstructive pulmonary disease, history of acute myocardial infarction, ischemic heart disease, diabetes, metastatic cancer and acute leukemia, history of stroke, depression, dementia, atrial fibrillation, rheumatoid arthritis or osteoarthritis, and chronic kidney disease.

Table 5.A.1.12b. ***(continued)***

[c] For observations in the baseline year, hospitalizations are measured in 2015, the year before the start of the baseline year. For observations in the intervention period, hospitalizations are measured in 2016, the year before the start of Program Year 1.

*/**/*** Significantly different from zero at the 0.10/0.05/0.01 level, two-tailed test.

HCC = hierarchical condition category; PY = Program Year; SSP = Medicare Shared Savings Program.

Table 5.A.1.13a. Regression-adjusted means and estimated impact of CPC+ on telehealth outcomes for attributed Medicare FFS beneficiaries in PYs 4 and 5, Track 1

	Track 1 – Overall				
	CPC+ mean[a]	C mean[a]	Impact estimate[b] (SE)	90 percent confidence interval	*p*-Value
Primary care visits					
Proportion of ambulatory primary care visits that are non-face-to- face[c,d]					
PY 4	15.7%	14.8%	0.9*** (0.3)	(0.3, 1.4)	0.01
PY 5	8.4%	8.1%	0.3* (0.2)	(0.0, 0.7)	0.08
Proportion of expenditures on ambulatory primary care visits that are non-face-to-face[c,d]					
PY 4	14.3%	13.7%	0.6* (0.4)	(0.0, 1.2)	0.08
PY 5	6.7%	6.5%	0.2 (0.2)	(-0.1, 0.5)	0.29
Unweighted sample sizes for non-face-to-face primary care visits proportion measure (PY 4 only)					
Number of practices	1,373	5,242			
Number of beneficiaries	921,251	3,204,785			
Unweighted sample sizes for non-face-to-face primary care visits proportion measure (PY 5 only)					
Number of practices	1,373	5,242			
Number of beneficiaries	932,199	3,299,749			
Unweighted sample sizes for non-face-to-face primary care expenditures proportion measure (PY 4 only)					
Number of practices	1,373	5,242			
Number of beneficiaries	873,540	3,026,256			
Unweighted sample sizes for non-face-to-face primary care expenditures proportion measure (PY 5 only)					
Number of practices	1,373	5,242			
Number of beneficiaries	899,255	3,171,076			

Source: Mathematica's analysis of Medicare claims data from January 2018 through December 2021.

Notes: Although this table indicates which estimates are statistically significant, when we interpret evidence, we combine evidence from the magnitude of the effect, the *p*-values, findings on related outcomes, subgroups, sensitivity tests, and other data sources about model implementation.

[a] The comparison group mean is computed by subtracting the regression adjusted difference between the CPC+ and comparison means in the year (PY 4 or PY 5) from the CPC+ mean in that year.

[b] Because non-face-to-face visits were close to zero in the baseline period (and the first three intervention years) for both CPC+ and comparison practices, we use a straight differences model for the non-face-to-face visit and expenditure outcomes. The estimate reflects the difference of the average outcome for Medicare FFS beneficiaries attributed to CPC+ practices in PY 4 or PY 5 to the average outcome for Medicare FFS beneficiaries attributed to comparison practices in the same time period while controlling for beneficiary characteristics and (selected) outcomes at baseline.

[c] Ambulatory visits are identified as face-to-face or non-face-to-face based on procedure codes, telehealth modifiers, and place of service (carrier file only) on Medicare claims. Visits such as telephone and online assessment and management and E&M are included in the non-face-to-face measure, making it broader than CMS's definition of "telehealth" visits.

[d] Measures include only beneficiaries with non-zero counts of visits or expenditures. Sample sizes for each measure are shown in table.

*/**/*** Significantly different from zero at the 0.10/0.05/0.01 level, two-tailed test.

C = comparison; CMS = Centers for Medicare and Medicaid Services; E&M = evaluation and management; FFS = fee-for-service; PY = Program Year; SE = standard error.

Table 5.A.1.13b. Regression-adjusted means and estimated impact of CPC+ on telehealth outcomes for attributed Medicare FFS beneficiaries in PYs 4 and 5, Track 1 by SSP status

	Track 1 – SSP					Track 1 – Non-SSP					
	CPC+ mean[a]	C mean[a]	Impact estimate[b] (SE)	90 percent confidence interval	*p*-Value	CPC+ mean[a]	C mean[a]	Impact estimate[b] (SE)	90 percent confidence interval	*p*-Value	*p*-Value for SSP vs. non-SSP difference
Primary care visits											
Proportion of ambulatory primary care visits that are non-face-to- face[c,d]											
PY 4	16.2%	15.5%	0.7* (0.4)	(0.0, 1.4)	0.08	15.1%	13.5%	1.6*** (0.5)	(0.8, 2.4)	0.00	0.18
PY 5	8.6%	8.5%	0.1 (0.3)	(-0.3, 0.6)	0.66	8.2%	7.5%	0.8*** (0.3)	(0.3, 1.2)	0.01	0.10
Proportion of expenditures on ambulatory primary care visits that are non-face-to-face[c,d]											
PY 4	14.6%	14.3%	0.3 (0.5)	(-0.5, 1.0)	0.58	14.0%	12.4%	1.6*** (0.5)	(0.7, 2.4)	0.00	0.06
PY 5	6.8%	6.9%	-0.1 (0.3)	(-0.5, 0.3)	0.68	6.6%	6.0%	0.6** (0.2)	(0.2, 1.0)	0.02	0.05
Unweighted sample sizes for non-face-to-face primary care visits proportion measure (PY 4 only)											
Number of practices	738	2,979				635	2,263				
Number of beneficiaries	464,981	1,877,161				456,270	1,327,624				
Unweighted sample sizes for non-face-to-face primary care visits proportion measure (PY 5 only)											
Number of practices	738	2,979				635	2,263				
Number of beneficiaries	470,098	1,936,446				462,101	1,363,303				
Unweighted sample sizes for non-face-to-face primary care expenditures proportion measure (PY 4 only)											
Number of practices	738	2,979				635	2,263				
Number of beneficiaries	441,291	1,776,890				432,249	1,249,366				
Unweighted sample sizes for non-face-to-face primary care expenditures proportion measure (PY 5 only)											
Number of practices	738	2,979				635	2,263				
Number of beneficiaries	453,787	1,862,762				445,468	1,308,314				

Source: Mathematica's analysis of Medicare claims data from January 2018 through December 2021.

Notes: Although this table indicates which estimates are statistically significant, when we interpret evidence, we combine evidence from the magnitude of the effect, the *p*-values, findings on related outcomes, subgroups, sensitivity tests, and other data sources about model implementation.

[a] The comparison group mean is computed by subtracting the regression adjusted difference between the CPC+ and comparison means in the year (PY 4 or PY 5) from the CPC+ mean in that year.

[b] Because non-face-to-face visits were close to zero in the baseline period (and the first three intervention years) for both CPC+ and comparison practices, we use a straight differences model for the non-face-to-face visit and expenditure outcomes. The estimate reflects the difference of the average outcome for Medicare FFS beneficiaries attributed to CPC+ practices in PY 4 or PY 5 to the average outcome for Medicare FFS beneficiaries attributed to comparison practices in the same time period while controlling for beneficiary characteristics and (selected) outcomes at baseline.

[c] Ambulatory visits are identified as face-to-face or non-face-to-face based on procedure codes, telehealth modifiers, and place of service (carrier file only) on Medicare claims. Visits such as telephone and online assessment and management and E&M are included in the non-face-to-face measure, making it broader than CMS's definition of "telehealth" visits.

[d] Measures include only beneficiaries with non-zero counts of visits or expenditures. Sample sizes for each measure are shown in table.

*/**/*** Significantly different from zero at the 0.10/0.05/0.01 level, two-tailed test.

C = comparison; CMS = Centers for Medicare and Medicaid Services; E&M = evaluation and management; FFS = fee-for-service; PY = Program Year; SE = standard error; SSP = Medicare Shared Savings Program.

Table 5.A.1.14a. Regression-adjusted means and estimated impact of CPC+ on telehealth outcomes for attributed Medicare FFS beneficiaries in PYs 4 and 5, Track 2

	Track 2 – Overall				
	CPC+ mean[a]	C mean[a]	Impact estimate[b] (SE)	90 percent confidence interval	*p*-Value
Primary care visits					
Proportion of ambulatory primary care visits that are non-face-to- face[c,d]					
PY 4	16.9%	14.8%	2.2*** (0.4)	(1.6, 2.7)	0.00
PY 5	8.9%	7.8%	1.1*** (0.2)	(0.8, 1.4)	0.00
Proportion of expenditures on ambulatory primary care visits that are non-face-to-face[c,d]					
PY 4	15.0%	13.6%	1.5*** (0.4)	(0.8, 2.1)	0.00
PY 5	6.8%	6.4%	0.5*** (0.2)	(0.2, 0.8)	0.01
Unweighted sample sizes for non-face-to-face primary care visits proportion measure (PY 4 only)					
Number of practices	1,515	3,783			
Number of beneficiaries	1,133,968	2,716,731			
Unweighted sample sizes for non-face-to-face primary care visits proportion measure (PY 5 only)					
Number of practices	1,515	3,783			
Number of beneficiaries	1,151,496	2,791,808			
Unweighted sample sizes for non-face-to-face primary care expenditures proportion measure (PY 4 only)					
Number of practices	1,515	3,783			
Number of beneficiaries	1,071,428	2,565,410			
Unweighted sample sizes for non-face-to-face primary care expenditures proportion measure (PY 5 only)					
Number of practices	1,515	3,783			
Number of beneficiaries	1,109,147	2,684,435			

Source: Mathematica's analysis of Medicare claims data from January 2018 through December 2021.

Notes: Although this table indicates which estimates are statistically significant, when we interpret evidence, we combine evidence from the magnitude of the effect, the *p*-values, findings on related outcomes, subgroups, sensitivity tests, and other data sources about model implementation.

[a] The comparison group mean is computed by subtracting the regression adjusted difference between the CPC+ and comparison means in the year (PY 4 or PY 5) from the CPC+ mean in that year.

[b] Because non-face-to-face visits were close to zero in the baseline period (and the first three intervention years) for both CPC+ and comparison practices, we use a straight differences model for the non-face-to-face visit and expenditure outcomes. The estimate reflects the difference of the average outcome for Medicare FFS beneficiaries attributed to CPC+ practices in PY 4 or PY 5 to the average outcome for Medicare FFS beneficiaries attributed to comparison practices in the same time period while controlling for beneficiary characteristics and (selected) outcomes at baseline.

[c] Ambulatory visits are identified as face-to-face or non-face-to-face based on procedure codes, telehealth modifiers, and place of service (carrier file only) on Medicare claims. Visits such as telephone and online assessment and management and E&M are included in the non-face-to-face measure, making it broader than CMS's definition of "telehealth" visits.

[d] Measures include only beneficiaries with non-zero counts of visits or expenditures. Sample sizes for each measure are shown in table.

*/**/*** Significantly different from zero at the 0.10/0.05/0.01 level, two-tailed test.

C = comparison; CMS = Centers for Medicare and Medicaid Services; E&M = evaluation and management; FFS = fee-for-service; PY = Program Year; SE = standard error.

Table 5.A.1.14b. Regression-adjusted means and estimated impact of CPC+ on telehealth outcomes for attributed Medicare FFS beneficiaries in PYs 4 and 5, Track 2 by SSP status

	Track 2 – SSP					Track 2 – Non-SSP					
	CPC+ mean[a]	C mean[a]	Impact estimate[b] (SE)	90 percent confidence interval	*p*-Value	CPC+ mean[a]	C mean[a]	Impact estimate[b] (SE)	90 percent confidence interval	*p*-Value	*p*-Value for SSP vs. non-SSP difference
Primary care visits											
Proportion of ambulatory primary care visits that are non-face-to- face[c,d]											
PY 4	17.7%	15.9%	1.9*** (0.6)	(0.9, 2.8)	0.00	16.3%	14.4%	1.9*** (0.4)	(1.2, 2.6)	0.00	0.93
PY 5	9.0%	7.9%	1.1*** (0.3)	(0.6, 1.7)	0.00	8.9%	7.9%	1.0*** (0.3)	(0.6, 1.4)	0.00	0.78
Proportion of expenditures on ambulatory primary care visits that are non-face-to-face[c,d]											
PY 4	15.5%	14.5%	1.0 (0.6)	(0.0, 1.9)	0.11	14.7%	13.4%	1.3*** (0.4)	(0.5, 2.0)	0.00	0.64
PY 5	6.9%	6.3%	0.5* (0.3)	(0.1, 1.0)	0.06	6.8%	6.4%	0.4* (0.3)	(0.0, 0.8)	0.10	0.79
Unweighted sample sizes for non-face-to-face primary care visits proportion measure (PY4 only)											
Number of practices	636	1,817				879	1,966				
Number of beneficiaries	500,936	1,362,631				633,032	1,354,100				
Unweighted sample sizes for non-face-to-face primary care visits proportion measure (PY5 only)											
Number of practices	636	1,817				879	1,966				
Number of beneficiaries	511,090	1,400,390				640,406	1,391,418				
Unweighted sample sizes for non-face-to-face primary care expenditures proportion measure (PY4 only)											
Number of practices	636	1,817				879	1,966				
Number of beneficiaries	474,199	1,290,923				597,229	1,274,487				
Unweighted sample sizes for non-face-to-face primary care expenditures proportion measure (PY5 only)											
Number of practices	636	1,817				879	1,966				
Number of beneficiaries	493,511	1,349,182				615,636	1,335,253				

Source: Mathematica's analysis of Medicare claims data from January 2018 through December 2021.

Notes: Although this table indicates which estimates are statistically significant, when we interpret evidence, we combine evidence from the magnitude of the effect, the *p*-values, findings on related outcomes, subgroups, sensitivity tests, and other data sources about model implementation.

[a] The comparison group mean is computed by subtracting the regression adjusted difference between the CPC+ and comparison means in the year (PY 4 or PY 5) from the CPC+ mean in that year.

[b] Because non-face-to-face visits were close to zero in the baseline period (and the first three intervention years) for both CPC+ and comparison practices, we use a straight differences model for the non-face-to-face visit and expenditure outcomes. The estimate reflects the difference of the average outcome for Medicare FFS beneficiaries attributed to CPC+ practices in PY 4 or PY 5 to the average outcome for Medicare FFS beneficiaries attributed to comparison practices in the same time period while controlling for beneficiary characteristics and (selected) outcomes at baseline.

[c] Ambulatory visits are identified as face-to-face or non-face-to-face based on procedure codes, telehealth modifiers, and place of service (carrier file only) on Medicare claims. Visits such as telephone and online assessment and management and E&M are included in the non-face-to-face measure, making it broader than CMS's definition of "telehealth" visits.

[d] Measures include only beneficiaries with non-zero counts of visits or expenditures. Sample sizes for each measure are shown in table.

*/**/*** Significantly different from zero at the 0.10/0.05/0.01 level, two-tailed test.

C = comparison; CMS = Centers for Medicare and Medicaid Services; E&M = evaluation and management; FFS = fee-for-service; PY = Program Year; SE = standard error; SSP = Medicare Shared Savings Program.

Table 5.A.1.15a. Estimated average annual impacts of CPC+ on outpatient ED visits across the five program years, from main analysis and sensitivity tests, Track 1

Test	Motivation	Track 1 – Overall				
		Impact estimate	Percentage impact	*p*-Value	90% CI lower bound	90% CI upper bound
Main analysis (average annual estimate for PY 1 through PY 5)	**Uses a difference-in-differences analysis with an ITT beneficiary sample, a one-year baseline period, controls for baseline beneficiary characteristics, COVID-19-related controls, and practice fixed effects**	**-9.4*** **	**-2.1%**	**0.00**	**-13.5**	**-5.3**
Altering length of baseline period						
Use two-year baseline period (instead of one year) [a]	Controls for outcome levels over longer pre-CPC+ period	-7.5***	-1.6%	0.00	-11.4	-3.6
Altering the composition of the beneficiary sample						
Use sample of beneficiaries attributed during both the baseline and intervention periods as the analysis sample [b]	Helps to adjust for changes in sample composition between baseline and follow-up that may differ for the intervention and matched comparison groups	-6.0***	-1.3%	0.01	-9.8	-2.2
Examine the impacts for the subset of beneficiaries attributed in the first quarter of the baseline period and the intervention period [c]	Removes any effects that may be due to changes in sample composition over time, for both baseline and intervention years	-6.4**	-1.4%	0.01	-10.5	-2.3
Instead of following an ITT approach to defining the beneficiary sample (once attributed, beneficiaries stay in the sample for all subsequent years), allow beneficiaries to drop out of the sample if they no longer meet attribution requirements [d, e]	Assesses whether ITT tends to attenuate true effects by retaining beneficiaries in the intervention group who are no longer seen by CPC+ practices	-10.4***	-2.4%	0.00	-14.5	-6.4
Altering the modeling assumptions						
Use baseline beneficiary characteristics, practice characteristics, and practice-level averages of beneficiary characteristics (reflecting baseline characteristics of contemporaneous beneficiaries), all interacted with year indicators as additional controls (confounder test)	Accounts for potential time-varying effects of baseline beneficiary and practice characteristics on the outcome. Adjusts for practice-level measures of beneficiary characteristics to align with participation in CPC+ varying at the practice level	-7.2***	-1.6%	0.00	-12.1	-2.3
Controlling for contemporaneous SSP participation						
Use a model that controls for contemporaneous (same year) SSP participation status	Controls for changes in SSP participation status among CPC+ and comparison practices over time	-9.9***	-2.2%	0.00	-14.0	-5.8
Alternative definition of counterfactual						
Use a triple differences approach [f]	Controls for regional differences in trends among CPC+ and comparison practices	-2.8	-0.6%	0.47	-9.2	3.6

Source: Mathematica's analysis of Medicare claims data from January 2013 through December 2021.

[a] Sample size is 14 percent larger than the main analysis.

[b] Sample size is 35 percent smaller than the main analysis.

[c] Sample size is 32 percent smaller than the main analysis.

Table 5.A.1.15a. ***(continued)***

[d] Sample size is 11 percent smaller than the main analysis.

[e] The percentage of beneficiaries that are no longer attributed to CPC+ or comparison practices but are still included in the research sample due to the ITT approach grows over time; however, the yearly estimate from this sensitivity test was similar to the corresponding estimate from the main analysis in PY 5 (-15.7 [$p < 0.01$] and -15.7 [$p < 0.01$], respectively).

[f] Sample size is 224 percent larger than the main analysis (because the triple-differences model also includes non-participating practices in CPC+ regions and unselected practices in comparison regions).

*/**/*** Significantly different from zero at the 0.10/0.05/0.01 level, two-tailed test.

CI = confidence interval; ED = emergency department; ITT = intent-to-treat; PY = Program Year; SSP = Medicare Shared Savings Program.

Table 5.A.1.15b. Estimated average annual impacts of CPC+ on outpatient ED visits across the five program years, from main analysis and sensitivity tests, Track 1 by SSP status

		Track 1 – SSP					Track 1 – Non-SSP				
Test	Motivation	Impact estimate	Percentage impact	*p*-Value	90% CI lower bound	90% CI upper bound	Impact estimate	Percentage impact	*p*-Value	90% CI lower bound	90% CI upper bound
Main analysis (average annual estimate for PY 1 through PY 5)	**Uses a difference-in-differences analysis with an ITT beneficiary sample, a one-year baseline period, controls for baseline beneficiary characteristics, COVID-19-related controls, and practice fixed effects**	**-10.2*** **	**-2.3%**	**0.00**	**-15.6**	**-4.8**	**-8.2****	**-1.8%**	**0.03**	**-14.4**	**-2.0**
Altering length of baseline period											
Use two-year baseline period (instead of one year) [a]	Controls for outcome levels over longer pre-CPC+ period	-8.6***	-2.0%	0.01	-13.7	-3.5	-6.0*	-1.3%	0.09	-12.0	-0.1
Altering the composition of the beneficiary sample											
Use sample of beneficiaries attributed during both the baseline and intervention periods as the analysis sample [b]	Helps to adjust for changes in sample composition between baseline and follow-up that may differ for the intervention and matched comparison groups	-6.8**	-1.5%	0.02	-11.8	-1.9	-4.9	-1.0%	0.16	-10.7	0.8
Examine the impacts for the subset of beneficiaries attributed in the first quarter of the baseline period and the intervention period [c]	Removes any effects that may be due to changes in sample composition over time, for both baseline and intervention years	-6.5**	-1.4%	0.04	-11.9	-1.2	-5.7	-1.2%	0.14	-12.0	0.6
Instead of following an ITT approach to defining the beneficiary sample (once attributed, beneficiaries stay in the sample for all subsequent years), allow beneficiaries to drop out of the sample if they no longer meet attribution requirements [d, e]	Assesses whether ITT tends to attenuate true effects by retaining beneficiaries in the intervention group who are no longer seen by CPC+ practices	-11.5***	-2.7%	0.00	-16.8	-6.2	-9.0**	-2.0%	0.02	-15.1	-2.9

Table 5.A.1.15b. ***(continued)***

Test	Motivation	Track 1 – SSP					Track 1 – Non-SSP				
		Impact estimate	Percentage impact	*p*-Value	90% CI lower bound	90% CI upper bound	Impact estimate	Percentage impact	*p*-Value	90% CI lower bound	90% CI upper bound
Altering the modeling assumptions											
Use baseline beneficiary characteristics, practice characteristics, and practice-level averages of beneficiary characteristics (reflecting baseline characteristics of contemporaneous beneficiaries), all interacted with year indicators as additional controls (confounder test)	Accounts for potential time-varying effects of baseline beneficiary and practice characteristics on the outcome. Adjusts for practice-level measures of beneficiary characteristics to align with participation in CPC+ varying at the practice level	-8.3***	-1.9%	0.01	-14.4	-2.2	-6.5*	-1.4%	0.09	-14.0	1.0
Controlling for contemporaneous SSP participation											
Use a model that controls for contemporaneous (same year) SSP participation status	Controls for changes in SSP participation status among CPC+ and comparison practices over time	-9.9***	-2.2%	0.00	-15.3	-4.5	-6.0	-1.3%	0.12	-12.2	0.3
Alternative definition of counterfactual											
Use a triple differences approach [f]	Controls for regional differences in trends among CPC+ and comparison practices	-4.3	-1.0%	0.45	-13.9	5.2	-0.7	-0.1%	0.89	-9.0	7.6

Source: Mathematica's analysis of Medicare claims data from January 2013 through December 2021.

[a] Sample size is 14 percent larger than the main analysis.

[b] Sample size is about 35 percent smaller than the main analysis.

[c] Sample size is 32 percent smaller than the main analysis.

[d] Sample size is about 11 percent smaller than the main analysis.

[e] The percentage of beneficiaries that are no longer attributed to CPC+ or comparison practices but are still included in the research sample due to the ITT approach grows over time; however, the yearly estimate from this sensitivity test was similar to the corresponding estimate from the main analysis in PY 5 (-18.2 [$p < 0.01$] and -17.6 [$p < 0.01$] for Track 1 SSP and -13.6 [$p = 0.02$] and -14.0 [$p = 0.01$] for Track 1 Non-SSP, respectively).

[f] Sample size is 129 to 348 percent larger than the main analysis (because the triple-differences model also includes non-participating practices in CPC+ regions and unselected practices in comparison regions).

*/**/*** Significantly different from zero at the 0.10/0.05/0.01 level, two-tailed test.

CI = confidence interval; ED = emergency department; ITT = intent-to-treat; PY = Program Year; SSP = Medicare Shared Savings Program.

Table 5.A.1.16a. Estimated average annual impacts of CPC+ on outpatient ED visits across the five program years, from main analysis and sensitivity tests, Track 2

Test	Motivation	Track 2 – Overall				
		Impact estimate	Percentage impact	*p*-Value	90% CI lower bound	90% CI upper bound
Main analysis (average annual estimate for PY 1 through PY 5)	**Uses a difference-in-differences analysis with an ITT beneficiary sample, a one-year baseline period, controls for baseline beneficiary characteristics, COVID-19-related controls, and practice fixed effects**	**-8.3*** **	**-1.8%**	**0.00**	**-12.6**	**-3.9**
Altering length of baseline period						
Use two-year baseline period (instead of one year) [a]	Controls for outcome levels over longer pre-CPC+ period	-8.1***	-1.8%	0.00	-12.3	-3.8
Altering the composition of the beneficiary sample						
Use sample of beneficiaries attributed during both the baseline and intervention periods as the analysis sample [b]	Helps to adjust for changes in sample composition between baseline and follow-up that may differ for the intervention and matched comparison groups	-6.3***	-1.3%	0.01	-10.1	-2.4
Examine the impacts for the subset of beneficiaries attributed in the first quarter of the baseline period and the intervention period [c]	Removes any effects that may be due to changes in sample composition over time, for both baseline and intervention years	-6.4**	-1.4%	0.01	-10.6	-2.2
Instead of following an ITT approach to defining the beneficiary sample (once attributed, beneficiaries stay in the sample for all subsequent years), allow beneficiaries to drop out of the sample if they no longer meet attribution requirements [d, e]	Assesses whether ITT tends to attenuate true effects by retaining beneficiaries in the intervention group who are no longer seen by CPC+ practices	-7.3***	-1.7%	0.00	-11.5	-3.2
Altering the modeling assumptions						
Use baseline beneficiary characteristics, practice characteristics, and practice-level averages of beneficiary characteristics (reflecting baseline characteristics of contemporaneous beneficiaries), all interacted with year indicators as additional controls (confounder test)	Accounts for potential time-varying effects of baseline beneficiary and practice characteristics on the outcome. Adjusts for practice-level measures of beneficiary characteristics to align with participation in CPC+ varying at the practice level	-3.0	-0.7%	0.23	-8.0	1.9
Controlling for contemporaneous SSP participation						
Use a model that controls for contemporaneous (same year) SSP participation status	Controls for changes in SSP participation status among CPC+ and comparison practices over time	-8.4***	-1.8%	0.00	-12.7	-4.0
Alternative definition of counterfactual						
Use a triple differences approach [f]	Controls for regional differences in trends among CPC+ and comparison practices	-0.2	0.0%	0.97	-7.3	7.0

Source: Mathematica's analysis of Medicare claims data from January 2013 through December 2021.

[a] Sample size is 14 percent larger than the main analysis.

[b] Sample size is 34 percent smaller than the main analysis.

[c] Sample size is 32 percent smaller than the main analysis.

[d] Sample size is 11 percent smaller than the main analysis.

Table 5.A.1.16a. ***(continued)***

[e] The percentage of beneficiaries that are no longer attributed to CPC+ or comparison practices but are still included in the research sample due to the ITT approach grows over time; however, the yearly estimate from this sensitivity test was similar to the corresponding estimate from the main analysis in PY 5 (-11.9 [p <0.01] and -11.2 [p <0.01], respectively).

[f] Sample size is 225 percent larger than the main analysis (because the triple-differences model also includes non-participating practices in CPC+ regions and unselected practices in comparison regions).

*/**/*** Significantly different from zero at the 0.10/0.05/0.01 level, two-tailed test.

CI = confidence interval; ITT = intent-to-treat; PY = Program Year.

Table 5.A.1.16b. Estimated average annual impacts of CPC+ on outpatient ED visits across the five program years, from main analysis and sensitivity tests, Track 2 by SSP status

		Track 2 – SSP					Track 2 – Non-SSP				
Test	Motivation	Impact estimate	Percentage impact	*p*-Value	90% CI lower bound	90% CI upper bound	Impact estimate	Percentage impact	*p*-Value	90% CI lower bound	90% CI upper bound
Main analysis (average annual estimate for PY 1 through PY 5)	**Uses a difference-in-differences analysis with an ITT beneficiary sample, a one-year baseline period, controls for baseline beneficiary characteristics, COVID-19-related controls, and practice fixed effects**	**-12.7*** **	**-2.9%**	**0.00**	**-19.0**	**-6.4**	**-3.8**	**-0.8%**	**0.28**	**-9.6**	**1.9**
Altering length of baseline period											
Use two-year baseline period (instead of one year) [a]	Controls for outcome levels over longer pre-CPC+ period	-13.8***	-3.1%	0.00	-20.1	-7.5	-2.8	-0.6%	0.41	-8.3	2.8
Altering the composition of the beneficiary sample											
Use sample of beneficiaries attributed during both the baseline and intervention periods as the analysis sample [b]	Helps to adjust for changes in sample composition between baseline and follow-up that may differ for the intervention and matched comparison groups	-8.9***	-2.0%	0.01	-14.3	-3.4	-3.5	-0.7%	0.27	-8.8	1.7
Examine the impacts for the subset of beneficiaries attributed in the first quarter of the baseline period and the intervention period [c]	Removes any effects that may be due to changes in sample composition over time, for both baseline and intervention years	-10.5***	-2.3%	0.01	-16.7	-4.4	-2.4	-0.5%	0.48	-8.0	3.2
Instead of following an ITT approach to defining the beneficiary sample (once attributed, beneficiaries stay in the sample for all subsequent years), allow beneficiaries to drop out of the sample if they no longer meet attribution requirements [d, e]	Assesses whether ITT tends to attenuate true effects by retaining beneficiaries in the intervention group who are no longer seen by CPC+ practices	-11.4***	-2.7%	0.00	-17.4	-5.4	-3.2	-0.7%	0.34	-8.8	2.3
Altering the modeling assumptions											
Use baseline beneficiary characteristics, practice characteristics, and practice-level averages of beneficiary characteristics (reflecting baseline characteristics of contemporaneous beneficiaries), all interacted with year indicators as additional controls (confounder test)	Accounts for potential time-varying effects of baseline beneficiary and practice characteristics on the outcome. Adjusts for practice-level measures of beneficiary characteristics to align with participation in CPC+ varying at the practice level	-5.1	-1.2%	0.19	-12.5	2.4	-0.5	-0.1%	0.87	-6.8	5.8

Table 5.A.1.16b. ***(continued)***

Test	Motivation	Track 2 – SSP					Track 2 – Non-SSP				
		Impact estimate	Percentage impact	*p*-Value	90% CI lower bound	90% CI upper bound	Impact estimate	Percentage impact	*p*-Value	90% CI lower bound	90% CI upper bound
Controlling for contemporaneous SSP participation											
Use a model that controls for contemporaneous (same year) SSP participation status	Controls for changes in SSP participation status among CPC+ and comparison practices over time	-12.1***	-2.7%	0.00	-18.5	-5.7	-2.2	-0.5%	0.54	-8.1	3.7
Alternative definition of counterfactual											
Use a triple differences approach [f]	Controls for regional differences in trends among CPC+ and comparison practices	-5.1	-1.2%	0.48	-17.0	6.8	4.3	0.9%	0.39	-3.8	12.3

Source: Mathematica's analysis of Medicare claims data from January 2013 through December 2021.

[a] Sample size is 14 percent larger than the main analysis.

[b] Sample size is 34 percent smaller than the main analysis.

[c] Sample size is 32 percent smaller than the main analysis.

[d] Sample size is about 11 percent smaller than the main analysis.

[e] The percentage of beneficiaries that are no longer attributed to CPC+ or comparison practices but are still included in the research sample due to the ITT approach grows over time; however, the yearly estimate from this sensitivity test was similar to the corresponding estimate from the main analysis in PY 5 (-21.0 [p <0.01] and -20.1 [p <0.01] for Track 2 SSP and -1.5 [p = 0.74] and -1.3 [p = 0.78] for Track 2 Non-SSP, respectively).

[f] Sample size is 155 to 290 percent larger than the main analysis (because the triple-differences model also includes non-participating practices in CPC+ regions and unselected practices in comparison regions).

*/**/*** Significantly different from zero at the 0.10/0.05/0.01 level, two-tailed test.

CI = confidence interval; ITT = intent-to-treat; PY = Program Year; SSP = Medicare Shared Savings Program.

Table 5.A.1.17a. Estimated impacts of CPC+ on outpatient ED visits in PYs 4 and 5, from main analysis and COVID-19 specific sensitivity tests, Track 1

	Track 1 – Overall				
Year	Impact estimate	Percentage impact	*p*-Value	90% CI lower bound	90% CI upper bound
Main analysis that uses a difference-in-differences empirical strategy					
PY 4 estimate	-10.8***	-2.8%	0.00	-16.3	-5.3
PY 5 estimate	-15.7***	-3.7%	0.00	-21.8	-9.7
Triple Differences Approach that controls for regional differences in trends due to COVID-19 among CPC+ and comparison practices[a]					
PY 4 estimate	-2.3	-0.6%	0.66	-11.0	6.4
PY 5 estimate	-6.7	-1.6%	0.24	-16.1	2.7
Estimates obtained for outcome constructed by dropping claims from March 2020 to May 2020 (to test for sensitivity to change in utilization at the start of the pandemic)[b]					
PY 4 estimate	-11.6***	-2.8%	0.00	-17.1	-6.0

Source: Mathematica's analysis of Medicare claims data from January 2013 through December 2021.

[a] Sample size is about 224 percent larger than the main analysis (because the triple-differences model also includes non-participating practices in CPC+ regions and unselected practices in comparison regions).

[b] Sample size is about 0.01 percent smaller than the main analysis.

*/**/*** Significantly different from zero at the 0.10/0.05/0.01 level, two-tailed test.

CI = confidence interval; ED = emergency department; PY = Program Year;; SSP = Medicare Shared Savings Program.

Table 5.A.1.17b. Estimated impacts of CPC+ on outpatient ED visits in PYs 4 and 5, from main analysis and COVID-19 specific sensitivity tests, Track 1 by SSP status

	Track 1 – SSP					Track 1 – Non-SSP				
Year	Impact estimate	Percentage impact	*p*-Value	90% CI lower bound	90% CI upper bound	Impact estimate	Percentage impact	*p*-Value	90% CI lower bound	90% CI upper bound
Main analysis that uses a difference-in-differences empirical strategy										
PY 4 estimate	-13.6***	-3.6%	0.00	-20.9	-6.3	-6.6	-1.7%	0.20	-15.0	1.9
PY 5 estimate	-17.6***	-4.3%	0.00	-25.4	-9.8	-14.0**	-3.2%	0.01	-23.0	-4.9
Triple Differences Approach that controls for regional differences in trends due to COVID-19 among CPC+ and comparison practices[a]										
PY 4 estimate	-4.0	-1.1%	0.62	-17.2	9.2	0.2	0.1%	0.97	-11.1	11.5
PY 5 estimate	-6.6	-1.7%	0.43	-20.6	7.3	-4.8	-1.1%	0.52	-17.2	7.5
Estimates obtained for outcome constructed by dropping claims from March 2020 to May 2020 (to test for sensitivity to change in utilization at the start of the pandemic)[b]										
PY 4 estimate	-13.7***	-3.5%	0.00	-21.1	-6.4	-8.2	-1.9%	0.11	-16.7	0.3

Source: Mathematica's analysis of Medicare claims data from January 2013 through December 2021.

[a] Sample size is 129 to 348 percent larger than the main analysis (because the triple-differences model also includes non-participating practices in CPC+ regions and unselected practices in comparison regions).

[b] Sample size is 0.01 percent smaller than the main analysis.

*/**/*** Significantly different from zero at the 0.10/0.05/0.01 level, two-tailed test.

CI = confidence interval; ED = emergency department; PY = Program Year; SSP = Medicare Shared Savings Program.

Table 5.A.1.18a. Estimated impacts of CPC+ on outpatient ED visits in PYs 4 and 5, from main analysis and COVID-19 specific sensitivity tests, Track 2

	Track 2 – Overall				
Year	Impact estimate	Percentage impact	*p*-Value	90% CI lower bound	90% CI upper bound
Main analysis that uses a difference-in-differences empirical strategy					
PY 4 estimate	-8.3**	-2.2%	0.02	-14.1	-2.5
PY 5 estimate	-11.2***	-2.7%	0.00	-17.2	-5.2
Triple Differences Approach that controls for regional differences in trends due to COVID-19 among CPC+ and comparison practices[a]					
PY 4 estimate	6.1	1.7%	0.26	-2.8	15.1
PY 5 estimate	3.5	0.9%	0.56	-6.4	13.4
Estimates obtained for outcome constructed by dropping claims from March 2020 to May 2020 (to test for sensitivity to change in utilization at the start of the pandemic)[b]					
PY 4 estimate	-8.6**	-2.1%	0.02	-14.4	-2.7

Source: Mathematica's analysis of Medicare claims data from January 2013 through December 2021.

[a] Sample size is 225 percent larger than the main analysis (because the triple-differences model also includes non-participating practices in CPC+ regions and unselected practices in comparison regions).

[b] Sample size is 0.01 percent smaller than the main analysis.

*/**/*** Significantly different from zero at the 0.10/0.05/0.01 level, two-tailed test.

CI = confidence interval; PY = Program Year.

Table 5.A.1.18b. Estimated impacts of CPC+ on outpatient ED visits in PYs 4 and 5, from main analysis and COVID-19 specific sensitivity tests, Track 2 by SSP status

	Track 2 – SSP					Track 2 – Non-SSP				
Year	Impact estimate	Percentage impact	*p*-Value	90% CI lower bound	90% CI upper bound	Impact estimate	Percentage impact	*p*-Value	90% CI lower bound	90% CI upper bound
Main analysis that uses a difference-in-differences empirical strategy										
PY 4 estimate	-19.2***	-5.0%	0.00	-27.9	-10.6	2.3	0.6%	0.63	-5.4	9.9
PY 5 estimate	-20.1***	-4.9%	0.00	-29.0	-11.3	-1.3	-0.3%	0.78	-8.9	6.4
Triple Differences Approach that controls for regional differences in trends due to COVID-19 among CPC+ and comparison practices[a]										
PY 4 estimate	-3.8	-1.0%	0.66	-17.8	10.3	15.4**	4.1%	0.02	4.6	26.1
PY 5 estimate	-2.4	-0.6%	0.81	-18.9	14.1	10.7	2.6%	0.12	-0.5	21.9
Estimates obtained for outcome constructed by dropping claims from March 2020 to May 2020 (to test for sensitivity to change in utilization at the start of the pandemic)[b]										
PY 4 estimate	-18.5***	-4.6%	0.00	-27.2	-9.8	1.4	0.3%	0.76	-6.3	9.2

Source: Mathematica's analysis of Medicare claims data from January 2013 through December 2021.

[a] Sample size is 155 to 290 percent larger than the main analysis (because the triple-differences model also includes non-participating practices in CPC+ regions and unselected practices in comparison regions).

[b] Sample size is 0.01 percent smaller than the main analysis.

*/**/*** Significantly different from zero at the 0.10/0.05/0.01 level, two-tailed test.

CI = confidence interval; PY = Program Year; SSP = Medicare Shared Savings Program.

Table 5.A.1.19a. Estimated average annual impacts of CPC+ on acute hospitalizations across the five program years, from main analysis and sensitivity tests, Track 1

Test	Motivation	Track 1 – Overall				
		Impact estimate	Percentage impact	*p*-Value	90% CI lower bound	90% CI upper bound
Main analysis (average annual estimate for PY 1 through PY 5)	**Uses a difference-in-differences analysis with an ITT beneficiary sample, a one-year baseline period, controls for baseline beneficiary characteristics, COVID-19-related controls, and practice fixed effects**	**-2.5***	**-0.9%**	**0.08**	**-4.9**	**-0.1**
Altering length of baseline period						
Use two-year baseline period (instead of one year) [a]	Controls for outcome levels over longer pre-CPC+ period	-2.2*	-0.8%	0.08	-4.4	-0.1
Altering the composition of the beneficiary sample						
Use sample of beneficiaries attributed during both the baseline and intervention periods as the analysis sample [b]	Helps to adjust for changes in sample composition between baseline and follow-up that may differ for the intervention and matched comparison groups	-2.0	-0.7%	0.17	-4.3	0.4
Examine the impacts for the subset of beneficiaries attributed in the first quarter of the baseline period and the intervention period [c]	Removes any effects that may be due to changes in sample composition over time, for both baseline and intervention years	-0.7	-0.3%	0.63	-3.2	1.8
Instead of following an ITT approach to defining the beneficiary sample (once attributed, beneficiaries stay in the sample for all subsequent years), allow beneficiaries to drop out of the sample if they no longer meet attribution requirements [d, e]	Assesses whether ITT tends to attenuate true effects by retaining beneficiaries in the intervention group who are no longer seen by CPC+ practices	-3.2**	-1.2%	0.03	-5.5	-0.8
Altering the modeling assumptions						
Use baseline beneficiary characteristics, practice characteristics, and practice-level averages of beneficiary characteristics (reflecting baseline characteristics of contemporaneous beneficiaries), all interacted with year indicators as additional controls (confounder test)	Accounts for potential time-varying effects of baseline beneficiary and practice characteristics on the outcome. Adjusts for practice-level measures of beneficiary characteristics to align with participation in CPC+ varying at the practice level	-0.3	-0.1%	0.82	-3.1	2.5
Controlling for contemporaneous SSP participation						
Use a model that controls for contemporaneous (same year) SSP participation status	Controls for changes in SSP participation status among CPC+ and comparison practices over time	-2.5*	-0.9%	0.09	-4.8	-0.1
Alternative definition of counterfactual						
Use a triple differences approach [f]	Controls for regional differences in trends among CPC+ and comparison practices	0.5	0.2%	0.80	-2.9	4.0

Source: Mathematica's analysis of Medicare claims data from January 2013 through December 2021.

[a] Sample size is 14 percent larger than the main analysis.

[b] Sample size is 35 percent smaller than the main analysis.

[c] Sample size is 32 percent smaller than the main analysis.

Table 5.A.1.19a. ***(continued)***

[d] Sample size is 11 percent smaller than the main analysis.

[e] The percentage of beneficiaries that are no longer attributed to CPC+ or comparison practices but are still included in the research sample due to the ITT approach grows over time; however, the yearly estimate from this sensitivity test was similar to the corresponding estimate from the main analysis in PY 5 (-3.5 [p = 0.06] and -2.5 [p = 0.15], respectively)..

[f] Sample size is 224 percent larger than the main analysis (because the triple-differences model also includes non-participating practices in CPC+ regions and unselected practices in comparison regions).

*/**/*** Significantly different from zero at the 0.10/0.05/0.01 level, two-tailed test.

CI = confidence interval; ITT = intent-to-treat; PY = Program Year.

Table 5.A.1.19b. Estimated average annual impacts of CPC+ on acute hospitalizations across the five program years, from main analysis and sensitivity tests, Track 1 by SSP status

		Track 1 – SSP					Track 1 – Non-SSP				
Test	Motivation	Impact estimate	Percentage impact	*p*-Value	90% CI lower bound	90% CI upper bound	Impact estimate	Percentage impact	*p*-Value	90% CI lower bound	90% CI upper bound
Main analysis (average annual estimate for PY 1 through PY 5)	**Uses a difference-in-differences analysis with an ITT beneficiary sample, a one-year baseline period, controls for baseline beneficiary characteristics, COVID-19-related controls, and practice fixed effects**	**-4.5****	**-1.6%**	**0.01**	**-7.5**	**-1.5**	**-0.1**	**0.0%**	**0.96**	**-3.8**	**3.6**
Altering length of baseline period											
Use two-year baseline period (instead of one year) [a]	Controls for outcome levels over longer pre-CPC+ period	-3.1*	-1.1%	0.06	-5.9	-0.4	-1.1	-0.4%	0.58	-4.4	2.2
Altering the composition of the beneficiary sample											
Use sample of beneficiaries attributed during both the baseline and intervention periods as the analysis sample [b]	Helps to adjust for changes in sample composition between baseline and follow-up that may differ for the intervention and matched comparison groups	-3.5*	-1.2%	0.07	-6.7	-0.3	-0.2	-0.1%	0.94	-3.7	3.4
Examine the impacts for the subset of beneficiaries attributed in the first quarter of the baseline period and the intervention period [c]	Removes any effects that may be due to changes in sample composition over time, for both baseline and intervention years	-2.0	-0.7%	0.31	-5.2	1.2	0.8	0.3%	0.75	-3.1	4.6
Instead of following an ITT approach to defining the beneficiary sample (once attributed, beneficiaries stay in the sample for all subsequent years), allow beneficiaries to drop out of the sample if they no longer meet attribution requirements [d, e]	Assesses whether ITT tends to attenuate true effects by retaining beneficiaries in the intervention group who are no longer seen by CPC+ practices	-4.8***	-1.8%	0.01	-7.8	-1.7	-1.1	-0.4%	0.63	-4.7	2.6

Table 5.A.1.19b. ***(continued)***

Test	Motivation	Track 1 – SSP					Track 1 – Non-SSP				
		Impact estimate	Percentage impact	*p*-Value	90% CI lower bound	90% CI upper bound	Impact estimate	Percentage impact	*p*-Value	90% CI lower bound	90% CI upper bound
Altering the modeling assumptions											
Use baseline beneficiary characteristics, practice characteristics, and practice-level averages of beneficiary characteristics (reflecting baseline characteristics of contemporaneous beneficiaries), all interacted with year indicators as additional controls (confounder test)	Accounts for potential time-varying effects of baseline beneficiary and practice characteristics on the outcome. Adjusts for practice-level measures of beneficiary characteristics to align with participation in CPC+ varying at the practice level	-2.5	-0.9%	0.17	-6.1	1.1	1.0	0.4%	0.62	-3.1	5.1
Controlling for contemporaneous SSP participation											
Use a model that controls for contemporaneous (same year) SSP participation status	Controls for changes in SSP participation status among CPC+ and comparison practices over time	-4.0**	-1.4%	0.03	-7.0	-0.9	0.3	0.1%	0.88	-3.3	4.0
Alternative definition of counterfactual											
Use a triple differences approach [f]	Controls for regional differences in trends among CPC+ and comparison practices	-0.2	-0.1%	0.95	-5.2	4.8	1.4	0.5%	0.62	-3.3	6.1

Source: Mathematica's analysis of Medicare claims data from January 2013 through December 2021.

[a] Sample size is 14 percent larger than the main analysis.

[b] Sample size is about 35 percent smaller than the main analysis.

[c] Sample size is 32 percent smaller than the main analysis.

[d] Sample size is about 11 percent smaller than the main analysis.

[e] The percentage of beneficiaries that are no longer attributed to CPC+ or comparison practices but are still included in the research sample due to the ITT approach grows over time; however, the yearly estimate from this sensitivity test was similar to the corresponding estimate from the main analysis in PY 5 (-4.8 [$p = 0.07$] and 5.1 [$p = 0.04$] for Track 1 SSP and -1.1 [$p = 0.70$] and 0.7 [$p = 0.81$] for Track 1 Non-SSP, respectively).

[f] Sample size is 129 to 348 percent larger than the main analysis (because the triple-differences model also includes non-participating practices in CPC+ regions and unselected practices in comparison regions).

*/**/*** Significantly different from zero at the 0.10/0.05/0.01 level, two-tailed test.

CI = confidence interval; ITT = intent-to-treat; PY = Program Year; SSP = Medicare Shared Savings Program.

Table 5.A.1.20a. Estimated average annual impacts of CPC+ on acute hospitalizations across the five program years, from main analysis and sensitivity tests, Track 2

Test	Motivation	Track 2 – Overall				
		Impact estimate	Percentage impact	*p*-Value	90% CI lower bound	90% CI upper bound
Main analysis (average annual estimate for PY 1 through PY 5)	**Uses a difference-in-differences analysis with an ITT beneficiary sample, a one-year baseline period, controls for baseline beneficiary characteristics, COVID-19-related controls, and practice fixed effects**	**-2.7***	**-1.0%**	**0.08**	**-5.3**	**-0.2**
Altering length of baseline period						
Use two-year baseline period (instead of one year) [a]	Controls for outcome levels over longer pre-CPC+ period	-2.3*	-0.8%	0.09	-4.5	-0.1
Altering the composition of the beneficiary sample						
Use sample of beneficiaries attributed during both the baseline and intervention periods as the analysis sample [b]	Helps to adjust for changes in sample composition between baseline and follow-up that may differ for the intervention and matched comparison groups	-1.8	-0.6%	0.22	-4.3	0.6
Examine the impacts for the subset of beneficiaries attributed in the first quarter of the baseline period and the intervention period [c]	Removes any effects that may be due to changes in sample composition over time, for both baseline and intervention years	-1.9	-0.7%	0.23	-4.6	0.7
Instead of following an ITT approach to defining the beneficiary sample (once attributed, beneficiaries stay in the sample for all subsequent years), allow beneficiaries to drop out of the sample if they no longer meet attribution requirements [d, e]	Assesses whether ITT tends to attenuate true effects by retaining beneficiaries in the intervention group who are no longer seen by CPC+ practices	-2.7*	-1.0%	0.08	-5.2	-0.2
Altering the modeling assumptions						
Use baseline beneficiary characteristics, practice characteristics, and practice-level averages of beneficiary characteristics (reflecting baseline characteristics of contemporaneous beneficiaries), all interacted with year indicators as additional controls (confounder test)	Accounts for potential time-varying effects of baseline beneficiary and practice characteristics on the outcome. Adjusts for practice-level measures of beneficiary characteristics to align with participation in CPC+ varying at the practice level	0.3	0.1%	0.85	-2.6	3.2
Controlling for contemporaneous SSP participation						
Use a model that controls for contemporaneous (same year) SSP participation status	Controls for changes in SSP participation status among CPC+ and comparison practices over time	-2.7*	-1.0%	0.08	-5.3	-0.2
Alternative definition of counterfactual						
Use a triple differences approach [f]	Controls for regional differences in trends among CPC+ and comparison practices	-1.0	-0.4%	0.65	-4.7	2.7

Source: Mathematica's analysis of Medicare claims data from January 2013 through December 2021.

[a] Sample size is 14 percent larger than the main analysis.

[b] Sample size is 34 percent smaller than the main analysis.

[c] Sample size is 32 percent smaller than the main analysis.

[d] Sample size is 11 percent smaller than the main analysis.

Table 5.A.1.20a. ***(continued)***

[e] The percentage of beneficiaries that are no longer attributed to CPC+ or comparison practices but are still included in the research sample due to the ITT approach grows over time; however, the yearly estimate from this sensitivity test was similar to the corresponding estimate from the main analysis in PY 5 (-2.6 [p = 0.20] and -1.9 [p = 0.34], respectively).

[f] Sample size is 225 percent larger than the main analysis (because the triple-differences model also includes non-participating practices in CPC+ regions and unselected practices in comparison regions).

*/**/*** Significantly different from zero at the 0.10/0.05/0.01 level, two-tailed test.

CI = confidence interval; ITT = intent-to-treat; PY = Program Year.

Table 5.A.1.20b. Estimated average annual impacts of CPC+ on acute hospitalizations across the five program years, from main analysis and sensitivity tests, Track 2 by SSP status

Test	Motivation	Track 2 – SSP					Track 2– Non-SSP				
		Impact estimate	Percentage impact	*p*-Value	90% CI lower bound	90% CI upper bound	Impact estimate	Percentage impact	*p*-Value	90% CI lower bound	90% CI upper bound
Main analysis (average annual estimate for PY 1 through PY 5)	**Uses a difference-in-differences analysis with an ITT beneficiary sample, a one-year baseline period, controls for baseline beneficiary characteristics, COVID-19-related controls, and practice fixed effects**	**-1.7**	**-0.6%**	**0.47**	**-5.7**	**2.2**	**-3.5***	**-1.3%**	**0.08**	**-6.8**	**-0.3**
Altering length of baseline period											
Use two-year baseline period (instead of one year) [a]	Controls for outcome levels over longer pre-CPC+ period	-1.6	-0.6%	0.45	-5.0	1.9	-2.8*	-1.1%	0.10	-5.6	0.0
Altering the composition of the beneficiary sample											
Use sample of beneficiaries attributed during both the baseline and intervention periods as the analysis sample [b]	Helps to adjust for changes in sample composition between baseline and follow-up that may differ for the intervention and matched comparison groups	0.1	0.0%	0.98	-3.7	3.9	-3.3*	-1.2%	0.08	-6.5	-0.2
Examine the impacts for the subset of beneficiaries attributed in the first quarter of the baseline period and the intervention period [c]	Removes any effects that may be due to changes in sample composition over time, for both baseline and intervention years	-0.6	-0.2%	0.81	-4.5	3.3	-3.0	-1.1%	0.16	-6.4	0.5
Instead of following an ITT approach to defining the beneficiary sample (once attributed, beneficiaries stay in the sample for all subsequent years), allow beneficiaries to drop out of the sample if they no longer meet attribution requirements [d, e]	Assesses whether ITT tends to attenuate true effects by retaining beneficiaries in the intervention group who are no longer seen by CPC+ practices	-1.1	-0.4%	0.65	-5.0	2.9	-3.9**	-1.5%	0.05	-7.1	-0.7
Altering the modeling assumptions											
Use baseline beneficiary characteristics, practice characteristics, and practice-level averages of beneficiary characteristics (reflecting baseline characteristics of contemporaneous beneficiaries), all interacted with year indicators as additional controls (confounder test)	Accounts for potential time-varying effects of baseline beneficiary and practice characteristics on the outcome. Adjusts for practice-level measures of beneficiary characteristics to align with participation in CPC+ varying at the practice level	1.1	0.4%	0.63	-3.4	5.5	-0.3	-0.1%	0.89	-4.1	3.6

Table 5.A.1.20b. ***(continued)***

Test	Motivation	Track 2 – SSP					Track 2– Non-SSP				
		Impact estimate	Percentage impact	*p*-Value	90% CI lower bound	90% CI upper bound	Impact estimate	Percentage impact	*p*-Value	90% CI lower bound	90% CI upper bound
Controlling for contemporaneous SSP participation											
Use a model that controls for contemporaneous (same year) SSP participation status	Controls for changes in SSP participation status among CPC+ and comparison practices over time	-1.1	-0.4%	0.63	-5.0	2.8	-3.2	-1.2%	0.11	-6.5	0.1
Alternative definition of counterfactual											
Use a triple differences approach [f]	Controls for regional differences in trends among CPC+ and comparison practices	-0.2	-0.1%	0.95	-6.3	5.8	-2.1	-0.8%	0.41	-6.4	2.2

Source: Mathematica's analysis of Medicare claims data from January 2013 through December 2021.

[a] Sample size is14 percent larger than the main analysis.

[b] Sample size is 34 percent smaller than the main analysis.

[c] Sample size is 32 percent smaller than the main analysis.

[d] Sample size is about 11 percent smaller than the main analysis.

[e] The percentage of beneficiaries that are no longer attributed to CPC+ or comparison practices but are still included in the research sample due to the ITT approach grows over time; however, the yearly estimate from this sensitivity test was similar to the corresponding estimate from the main analysis in PY 5 (-1.6 [$p = 0.63$] and -2.3 [$p = 0.46$] for Track 2 SSP and -3.5 [$p = 0.16$] and -1.9 [$p = 0.44$] for Track 2 Non-SSP, respectively).

[f] Sample size is 155 to 290 percent larger than the main analysis (because the triple-differences model also includes non-participating practices in CPC+ regions and unselected practices in comparison regions).

*/**/*** Significantly different from zero at the 0.10/0.05/0.01 level, two-tailed test.

CI = confidence interval; ITT = intent-to-treat; PY = Program Year; SSP = Medicare Shared Savings Program.

Table 5.A.1.21a. Estimated impacts of CPC+ on acute hospitalizations in PYs 4 and 5, from main analysis and COVID-19 specific sensitivity tests, Track 1

	Track 1 – Overall				
Year	Impact estimate	Percentage impact	*p*-Value	90% CI lower bound	90% CI upper bound
Main analysis that uses a difference-in-differences empirical strategy					
PY 4 estimate	-4.9***	-2.0%	0.01	-7.8	-2.0
PY 5 estimate	-2.6	-1.1%	0.15	-5.6	0.4
Triple Differences Approach that controls for regional differences in trends due to COVID-19 among CPC+ and comparison practices[a]					
PY 4 estimate	2.1	0.9%	0.43	-2.3	6.5
PY 5 estimate	0.2	0.1%	0.94	-4.3	4.7
Estimates obtained for outcome constructed by dropping claims from March 2020 to May 2020 (to test for sensitivity to change in utilization at the start of the pandemic)[b]					
PY 4 estimate	-4.2**	-1.6%	0.02	-7.2	-1.1

Source: Mathematica's analysis of Medicare claims data from January 2013 through December 2021.

[a] Sample size is 224 percent larger than the main analysis (because the triple-differences model also includes non-participating practices in CPC+ regions and unselected practices in comparison regions).

[b] Sample size is about 0.01 percent smaller than the main analysis.

*/**/*** Significantly different from zero at the 0.10/0.05/0.01 level, two-tailed test.

CI = confidence interval; PY = Program Year.

Table 5.A.1.21b. Estimated impacts of CPC+ on acute hospitalizations in PYs 4 and 5, from main analysis and COVID-19 specific sensitivity tests, Track 1 by SSP status

	Track 1 – SSP					Track 1 – Non-SSP				
Year	Impact estimate	Percentage impact	*p*-Value	90% CI lower bound	90% CI upper bound	Impact estimate	Percentage impact	*p*-Value	90% CI lower bound	90% CI upper bound
Main analysis that uses a difference-in-differences empirical strategy										
PY 4 estimate	-8.0***	-3.2%	0.00	-11.9	-4.2	-1.4	-0.6%	0.62	-5.9	3.2
PY 5 estimate	-5.1**	-2.0%	0.04	-9.2	-1.0	0.7	0.3%	0.81	-3.7	5.0
Triple Differences Approach that controls for regional differences in trends due to COVID-19 among CPC+ and comparison practices[a]										
PY 4 estimate	0.3	0.1%	0.94	-6.1	6.6	4.0	1.7%	0.27	-2.0	10.0
PY 5 estimate	-3.4	-1.3%	0.41	-10.1	3.3	4.5	1.9%	0.21	-1.4	10.5
Estimates obtained for outcome constructed by dropping claims from March 2020 to May 2020 (to test for sensitivity to change in utilization at the start of the pandemic)[b]										
PY 4 estimate	-7.6***	-2.8%	0.00	-11.5	-3.7	-0.4	-0.2%	0.89	-5.1	4.3

Source: Mathematica's analysis of Medicare claims data from January 2013 through December 2021.

[a] Sample size is 129 to 348 percent larger than the main analysis (because the triple-differences model also includes non-participating practices in CPC+ regions and unselected practices in comparison regions).

[b] Sample size is 0.01 percent smaller than the main analysis.

*/**/*** Significantly different from zero at the 0.10/0.05/0.01 level, two-tailed test.

CI = confidence interval; PY = Program Year; SSP = Medicare Shared Savings Program.

Table 5.A.1.22a. Estimated impacts of CPC+ on acute hospitalizations in PYs 4 and 5, from main analysis and COVID-19 specific sensitivity tests, Track 2

	Track 2 – Overall				
Year	Impact estimate	Percentage impact	*p*-Value	90% CI lower bound	90% CI upper bound
Main analysis that uses a difference-in-differences empirical strategy					
PY 4 estimate	-4.8**	-1.9%	0.01	-8.0	-1.7
PY 5 estimate	-1.9	-0.8%	0.34	-5.1	1.3
Triple Differences Approach that controls for regional differences in trends due to COVID-19 among CPC+ and comparison practices[a]					
PY 4 estimate	0.5	0.2%	0.87	-4.1	5.0
PY 5 estimate	-2.3	-0.9%	0.41	-7.0	2.4
Estimates obtained for outcome constructed by dropping claims from March 2020 to May 2020 (to test for sensitivity to change in utilization at the start of the pandemic)[b]					
PY 4 estimate	-4.7**	-1.8%	0.02	-7.9	-1.4

Source: Mathematica's analysis of Medicare claims data from January 2013 through December 2021.

[a] Sample size is 225 percent larger than the main analysis (because the triple-differences model also includes non-participating practices in CPC+ regions and unselected practices in comparison regions).

[b] Sample size is 0.01 percent smaller than the main analysis.

*/**/*** Significantly different from zero at the 0.10/0.05/0.01 level, two-tailed test.

CI = confidence interval; PY = Program Year.

Table 5.A.1.22b. Estimated impacts of CPC+ on acute hospitalizations in PYs 4 and 5, from main analysis and COVID-19 specific sensitivity tests, Track 2 by SSP status

	Track 2 – SSP					Track 2 – Non-SSP				
Year	Impact estimate	Percentage impact	*p*-Value	90% CI lower bound	90% CI upper bound	Impact estimate	Percentage impact	*p*-Value	90% CI lower bound	90% CI upper bound
Main analysis that uses a difference-in-differences empirical strategy										
PY 4 estimate	-3.9	-1.5%	0.19	-8.9	1.1	-5.1**	-2.1%	0.03	-8.9	-1.2
PY 5 estimate	-2.3	-0.9%	0.46	-7.5	2.9	-1.9	-0.8%	0.44	-5.9	2.1
Triple Differences Approach that controls for regional differences in trends due to COVID-19 among CPC+ and comparison practices[a]										
PY 4 estimate	3.0	1.2%	0.51	-4.4	10.3	-2.0	-0.8%	0.54	-7.3	3.3
PY 5 estimate	-5.8	-2.2%	0.23	-13.7	2.1	-0.6	-0.2%	0.86	-5.9	4.7
Estimates obtained for outcome constructed by dropping claims from March 2020 to May 2020 (to test for sensitivity to change in utilization at the start of the pandemic)[b]										
PY 4 estimate	-3.8	-1.4%	0.22	-9.0	1.3	-5.0**	-2.0%	0.04	-9.0	-1.1

Source: Mathematica's analysis of Medicare claims data from January 2013 through December 2021.

[a] Sample size is 155 to 290 percent larger than the main analysis (because the triple-differences model also includes non-participating practices in CPC+ regions and unselected practices in comparison regions).

[b] Sample size is 0.01 percent smaller than the main analysis.

*/**/*** Significantly different from zero at the 0.10/0.05/0.01 level, two-tailed test.

CI = confidence interval; PY = Program Year.

5.A.2. Medicare FFS expenditures

Table 5.A.2.1a. Regression-adjusted means and estimated impacts of CPC+ on selected Medicare expenditure outcomes for attributed Medicare FFS beneficiaries by program year and average across the five program years, Track 1

	Track 1—Overall					
	CPC+ mean[a]	C mean[a]	Impact estimate[b] (SE)	Percentage impact[c]	90 percent confidence interval	*p*-Value
Medicare expenditures (per beneficiary per month)						
Medicare Part A and B expenditures without enhanced payments for CPC+ and SSP[d]						
Baseline	$881	$884	NA	NA	NA	NA
PY 1	$899	$897	$4.9 ($3.4)	0.6%	(-$0.7, $10.5)	0.15
PY 2	$949	$948	$4.0 ($3.6)	0.4%	(-$1.9, $9.9)	0.27
PY 3	$994	$995	$2.2 ($4.1)	0.2%	(-$4.5, $9.0)	0.59
PY 4	$949	$955	-$2.8 ($4.5)	-0.3%	(-$10.2, $4.7)	0.54
PY 5	$1,042	$1,048	-$3.1 ($4.7)	-0.3%	(-$10.9, $4.6)	0.51
PY 1 through 5	$969	$971	$1.1 ($3.3)	0.1%	(-$4.3, $6.6)	0.74
Medicare Part A and B expenditures including care management fees						
Baseline	$881	$884	NA	NA	NA	NA
PY 1	$913	$897	$18.8*** ($3.4)	2.1%	($13.2, $24.3)	0.00
PY 2	$962	$948	$16.6*** ($3.6)	1.8%	($10.7, $22.5)	0.00
PY 3	$1,006	$995	$14.0*** ($4.1)	1.4%	($7.2, $20.7)	0.00
PY 4	$960	$955	$8.3* ($4.5)	0.9%	($0.8, $15.7)	0.07
PY 5	$1,052	$1,048	$6.9 ($4.7)	0.7%	(-$0.8, $14.7)	0.14
PY 1 through 5	$981	$971	$13.0*** ($3.3)	1.3%	($7.5, $18.4)	0.00
Medicare Part A and B expenditures including care management fees, Performance-based Incentive Payments, and shared savings payments to SSP ACOs[e]						
Baseline	$883	$886	NA	NA	NA	NA
PY 1	$917	$900	$19.6*** ($3.4)	2.2%	($14.0, $25.2)	0.00
PY 2	$966	$951	$17.8*** ($3.6)	1.9%	($11.9, $23.7)	0.00
PY 3	$1,011	$1,000	$14.5*** ($4.1)	1.5%	($7.9, $21.2)	0.00
PY 4	$966	$963	$6.2 ($4.5)	0.6%	(-$1.2, $13.6)	0.17
PY 5	$1,057	$1,054	$5.0 ($4.7)	0.5%	(-$2.7, $12.8)	0.28
PY 1 through 5	$986	$976	$12.8*** ($3.3)	1.3%	($7.4, $18.2)	0.00

Table 5.A.2.1a. ***(continued)***

	Track 1—Overall					
	CPC+ mean[a]	C mean[a]	Impact estimate[b] (SE)	Percentage impact[c]	90 percent confidence interval	*p*-Value
Medicare expenditures by service category (per beneficiary per month)[f]						
Inpatient expenditures						
Baseline	$311	$318	NA	NA	NA	$311
PY 1	$316	$320	$2.8 ($2.3)	0.9%	(-$0.9, $6.6)	$316
PY 2	$322	$328	$0.8 ($2.3)	0.2%	(-$3.0, $4.5)	$322
PY 3	$332	$342	-$2.4 ($2.5)	-0.7%	(-$6.5, $1.8)	$332
PY 4	$315	$328	-$5.2** ($2.6)	-1.6%	(-$9.5, -$0.9)	$315
PY 5	$330	$341	-$3.2 ($2.7)	-1.0%	(-$7.7, $1.2)	$330
PY 1 through 5	$323	$332	-$1.4 ($2.0)	-0.4%	(-$4.7, $1.9)	$323
Expenditures for acute inpatient care[g]						
Baseline	$275	$282	NA	NA	NA	$275
PY 1	$279	$285	$1.2 ($2.0)	0.4%	(-$2.1, $4.5)	$279
PY 2	$285	$293	-$1.5 ($2.0)	-0.5%	(-$4.9, $1.8)	$285
PY 3	$295	$306	-$4.4** ($2.2)	-1.5%	(-$8.0, -$0.8)	$295
PY 4	$280	$293	-$6.9*** ($2.3)	-2.4%	(-$10.7, -$3.1)	$280
PY 5	$294	$305	-$3.8 ($2.4)	-1.3%	(-$7.8, $0.2)	$294
PY 1 through 5	$287	$296	-$3.1* ($1.8)	-1.1%	(-$6.0, -$0.2)	$287
Expenditures for acute surgical hospitalizations						
Baseline	$148	$149	NA	NA	NA	NA
PY 1	$148	$147	$2.6* ($1.5)	1.8%	($0.1, $5.1)	0.09
PY 2	$149	$149	$0.4 ($1.5)	0.3%	(-$2.0, $2.8)	0.79
PY 3	$154	$155	-$0.5 ($1.6)	-0.3%	(-$3.1, $2.1)	0.74
PY 4	$138	$140	-$1.0 ($1.6)	-0.7%	(-$3.6, $1.5)	0.51
PY 5	$143	$145	-$1.0 ($1.7)	-0.7%	(-$3.7, $1.8)	0.56
PY 1 through 5	$146	$147	$0.1 ($1.2)	0.1%	(-$2.0, $2.1)	0.94
Expenditures for acute medical hospitalizations						
Baseline	$127	$133	NA	NA	NA	NA
PY 1	$131	$138	-$1.4 ($1.0)	-1.1%	(-$3.1, $0.3)	0.17
PY 2	$136	$144	-$1.9* ($1.1)	-1.4%	(-$3.8, -$0.1)	0.08
PY 3	$141	$151	-$3.9*** ($1.2)	-2.7%	(-$5.9, -$1.8)	0.00
PY 4	$141	$153	-$5.9*** ($1.4)	-4.0%	(-$8.1, -$3.6)	0.00
PY 5	$151	$160	-$2.8* ($1.4)	-1.8%	(-$5.2, -$0.4)	0.05
PY 1 through 5	$140	$149	-$3.2*** ($1.0)	-2.2%	(-$4.8, -$1.5)	0.00

Table 5.A.2.1a. ***(continued)***

	Track 1—Overall					
	CPC+ mean[a]	C mean[a]	Impact estimate[b] (SE)	Percentage impact[c]	90 percent confidence interval	*p*-Value
Inpatient rehabilitation facility expenditures						
Baseline	$20	$21	NA	NA	NA	NA
PY 1	$22	$21	$0.8* ($0.4)	3.9%	($0.1, $1.5)	0.07
PY 2	$23	$22	$1.6*** ($0.5)	7.3%	($0.8, $2.3)	0.00
PY 3	$24	$23	$1.1** ($0.5)	5.0%	($0.3, $2.0)	0.03
PY 4	$23	$22	$1.3** ($0.5)	5.9%	($0.4, $2.1)	0.01
PY 5	$24	$24	$0.7 ($0.5)	2.9%	(-$0.2, $1.6)	0.21
PY 1 through 5	$23	$23	$1.1*** ($0.4)	5.0%	($0.5, $1.7)	0.01
Post-acute care expenditures[h]						
Baseline	$110	$112	NA	NA	NA	NA
PY 1	$109	$109	$1.5 ($1.1)	1.4%	(-$0.4, $3.4)	0.19
PY 2	$107	$109	$1.0 ($1.1)	1.0%	(-$0.9, $2.9)	0.38
PY 3	$106	$109	-$0.3 ($1.3)	-0.3%	(-$2.4, $1.8)	0.81
PY 4	$95	$97	$0.4 ($1.4)	0.4%	(-$1.9, $2.6)	0.78
PY 5	$92	$95	-$0.2 ($1.3)	-0.2%	(-$2.4, $2.0)	0.89
PY 1 through 5	$101	$103	$0.5 ($1.0)	0.5%	(-$1.2, $2.1)	0.64
Acute inpatient and post-acute care expenditures combined[h]						
Baseline	$385	$394	NA	NA	NA	NA
PY 1	$388	$394	$2.7 ($2.7)	0.7%	(-$1.7, $7.2)	0.32
PY 2	$392	$402	-$0.5 ($2.7)	-0.1%	(-$4.9, $3.8)	0.84
PY 3	$401	$415	-$4.7 ($3.0)	-1.2%	(-$9.6, $0.2)	0.12
PY 4	$375	$390	-$6.5** ($3.2)	-1.7%	(-$11.7, -$1.3)	0.04
PY 5	$386	$399	-$4.0 ($3.3)	-1.0%	(-$9.3, $1.4)	0.22
PY 1 through 5	$388	$400	-$2.6 ($2.4)	-0.7%	(-$6.5, $1.3)	0.27
Outpatient expenditures						
Baseline	$165	$170	NA	NA	NA	NA
PY 1	$177	$180	$0.8 ($0.8)	0.5%	(-$0.6, $2.2)	0.34
PY 2	$199	$201	$1.7 ($1.1)	0.8%	(-$0.2, $3.5)	0.14
PY 3	$214	$217	$1.1 ($1.3)	0.5%	(-$1.1, $3.2)	0.41
PY 4	$204	$208	-$0.2 ($1.5)	-0.1%	(-$2.8, $2.3)	0.89
PY 5	$232	$240	-$3.7** ($1.7)	-1.6%	(-$6.5, -$0.9)	0.03
PY 1 through 5	$206	$210	$0.0 ($1.1)	0.0%	(-$1.8, $1.7)	0.98

Table 5.A.2.1a. ***(continued)***

	Track 1—Overall					
	CPC+ mean[a]	C mean[a]	Impact estimate[b] (SE)	Percentage impact[c]	90 percent confidence interval	*p*-Value
Expenditures for outpatient ED visits, including observation stays[i]						
Baseline	$25	$26	NA	NA	NA	NA
PY 1	$26	$27	$0.1 ($0.2)	0.3%	(-$0.2, $0.4)	0.63
PY 2	$28	$29	$0.0 ($0.2)	0.1%	(-$0.4, $0.4)	0.91
PY 3	$29	$30	-$0.1 ($0.2)	-0.3%	(-$0.5, $0.3)	0.73
PY 4	$24	$25	$0.0 ($0.2)	-0.1%	(-$0.4, $0.4)	0.90
PY 5	$28	$30	-$0.6* ($0.3)	-2.2%	(-$1.2, -$0.1)	0.06
PY 1 through 5	$27	$28	-$0.1 ($0.2)	-0.5%	(-$0.5, $0.2)	0.55
Expenditures for physician and nonphysician Part B noninstitutional services in any setting[j]						
Baseline	$254	$242	NA	NA	NA	NA
PY 1	$258	$247	$0.0 ($0.8)	0.0%	(-$1.3, $1.3)	0.99
PY 2	$275	$262	$1.3 ($1.0)	0.5%	(-$0.3, $3.0)	0.19
PY 3	$289	$275	$2.7** ($1.2)	0.9%	($0.7, $4.7)	0.03
PY 4	$271	$256	$3.3** ($1.4)	1.2%	($1.1, $5.6)	0.02
PY 5	$315	$301	$2.8* ($1.6)	0.9%	($0.2, $5.4)	0.08
PY 1 through 5	$283	$269	$2.1** ($1.0)	0.7%	($0.4, $3.7)	0.04
Expenditures for ambulatory visits with primary care practitioners						
Baseline	$24	$24	NA	NA	NA	NA
PY 1	$24	$25	-$0.2* ($0.1)	-0.7%	(-$0.3, $0.0)	0.09
PY 2	$25	$26	$0.0 ($0.1)	-0.2%	(-$0.3, $0.2)	0.77
PY 3	$27	$27	-$0.2 ($0.2)	-0.6%	(-$0.4, $0.1)	0.28
PY 4	$24	$25	-$0.3 ($0.2)	-1.3%	(-$0.6, $0.0)	0.10
PY 5	$31	$32	-$0.5** ($0.2)	-1.7%	(-$0.9, -$0.1)	0.03
PY 1 through 5	$26	$27	-$0.2* ($0.1)	-0.9%	(-$0.5, $0.0)	0.08
Expenditures for ambulatory visits with primary care practitioners at assigned practice[k]						
Baseline	$17	$17	NA	NA	NA	NA
PY 1	$17	$16	-$0.1 ($0.1)	-0.5%	(-$0.3, $0.1)	0.46
PY 2	$15	$15	$0.1 ($0.2)	1.0%	(-$0.1, $0.4)	0.35
PY 3	$15	$15	$0.1 ($0.2)	0.7%	(-$0.2, $0.4)	0.60
PY 4	$12	$12	-$0.1 ($0.2)	-0.6%	(-$0.5, $0.3)	0.77
PY 5	$14	$14	$0.0 ($0.3)	0.3%	(-$0.4, $0.5)	0.89
PY 1 through 5	$15	$14	$0.0 ($0.2)	0.2%	(-$0.2, $0.3)	0.85

Table 5.A.2.1a. ***(continued)***

	Track 1—Overall					
	CPC+ mean[a]	C mean[a]	Impact estimate[b] (SE)	Percentage impact[c]	90 percent confidence interval	*p*-Value
Expenditures for ambulatory visits with primary care practitioners at non-assigned practice[k]						
Baseline	$7	$7	NA	NA	NA	NA
PY 1	$8	$8	-$0.1 ($0.1)	-1.2%	(-$0.2, $0.0)	0.16
PY 2	$10	$11	-$0.2 ($0.1)	-1.8%	(-$0.4, $0.0)	0.10
PY 3	$11	$12	-$0.3* ($0.1)	-2.3%	(-$0.5, $0.0)	0.05
PY 4	$12	$13	-$0.2 ($0.2)	-2.0%	(-$0.5, $0.1)	0.19
PY 5	$16	$17	-$0.6** ($0.2)	-3.4%	(-$0.9, -$0.2)	0.01
PY 1 through 5	$12	$12	-$0.3** ($0.1)	-2.3%	(-$0.5, -$0.1)	0.03
Expenditures for ambulatory visits with specialists						
Baseline	$25	$24	NA	NA	NA	NA
PY 1	$25	$24	$0.1 ($0.1)	0.4%	($0.0, $0.2)	0.14
PY 2	$26	$24	$0.2** ($0.1)	0.7%	($0.0, $0.3)	0.03
PY 3	$26	$25	$0.1 ($0.1)	0.3%	(-$0.1, $0.2)	0.43
PY 4	$22	$21	$0.0 ($0.1)	0.0%	(-$0.2, $0.2)	0.99
PY 5	$30	$29	-$0.2 ($0.1)	-0.7%	(-$0.4, $0.0)	0.13
PY 1 through 5	$26	$25	$0.0 ($0.1)	0.1%	(-$0.1, $0.2)	0.70
Expenditures on laboratory services[l]						
Baseline	$24	$24	NA	NA	NA	NA
PY 1	$25	$25	$0.0 ($0.1)	0.0%	(-$0.2, $0.2)	0.97
PY 2	$27	$27	$0.2* ($0.1)	0.9%	($0.0, $0.5)	0.07
PY 3	$27	$27	$0.0 ($0.2)	0.2%	(-$0.3, $0.3)	0.81
PY 4	$29	$29	-$0.2 ($0.2)	-0.8%	(-$0.6, $0.1)	0.23
PY 5	$34	$34	-$0.7** ($0.3)	-2.1%	(-$1.3, -$0.2)	0.02
PY 1 through 5	$29	$29	-$0.1 ($0.1)	-0.4%	(-$0.4, $0.1)	0.38
Expenditures on imaging services[l]						
Baseline	$46	$46	NA	NA	NA	NA
PY 1	$47	$47	-$0.5** ($0.2)	-1.0%	(-$0.9, -$0.1)	0.03
PY 2	$51	$50	-$0.4 ($0.3)	-0.7%	(-$0.8, $0.1)	0.16
PY 3	$53	$53	-$0.2 ($0.3)	-0.4%	(-$0.7, $0.3)	0.51
PY 4	$48	$47	-$0.1 ($0.3)	-0.1%	(-$0.5, $0.4)	0.87
PY 5	$54	$54	-$0.8*** ($0.3)	-1.5%	(-$1.4, -$0.3)	0.01
PY 1 through 5	$51	$50	-$0.4* ($0.2)	-0.8%	(-$0.8, $0.0)	0.08

Table 5.A.2.1a. ***(continued)***

	Track 1—Overall					
	CPC+ mean[a]	C mean[a]	Impact estimate[b] (SE)	Percentage impact[c]	90 percent confidence interval	*p*-Value
Skilled nursing facility expenditures						
Baseline	$67	$68	NA	NA	NA	NA
PY 1	$65	$66	$0.4 ($0.7)	0.7%	(-$0.7, $1.6)	0.52
PY 2	$64	$66	-$0.1 ($0.7)	-0.2%	(-$1.3, $1.1)	0.89
PY 3	$63	$65	-$0.4 ($0.8)	-0.7%	(-$1.7, $0.8)	0.57
PY 4	$64	$67	-$1.6 ($1.2)	-2.5%	(-$3.6, $0.3)	0.18
PY 5	$65	$66	-$0.3 ($1.1)	-0.5%	(-$2.1, $1.5)	0.77
PY 1 through 5	$64	$66	-$0.4 ($0.7)	-0.6%	(-$1.5, $0.7)	0.56
Home health expenditures						
Baseline	$39	$41	NA	NA	NA	NA
PY 1	$39	$41	-$0.3 ($0.3)	-0.6%	(-$0.7, $0.2)	0.37
PY 2	$39	$42	-$1.0*** ($0.3)	-2.5%	(-$1.5, -$0.5)	0.00
PY 3	$39	$42	-$1.6*** ($0.4)	-3.8%	(-$2.1, -$1.0)	0.00
PY 4	$36	$39	-$1.7*** ($0.4)	-4.5%	(-$2.4, -$1.0)	0.00
PY 5	$38	$42	-$2.0*** ($0.4)	-5.1%	(-$2.8, -$1.3)	0.00
PY 1 through 5	$38	$41	-$1.3*** ($0.3)	-3.3%	(-$1.8, -$0.8)	0.00
Hospice expenditures						
Baseline	$23	$24	NA	NA	NA	NA
PY 1	$24	$24	$1.1*** ($0.4)	4.8%	($0.4, $1.8)	0.01
PY 2	$27	$27	$1.6*** ($0.5)	6.4%	($0.8, $2.4)	0.00
PY 3	$31	$30	$2.4*** ($0.6)	8.6%	($1.5, $3.4)	0.00
PY 4	$32	$32	$2.2*** ($0.6)	7.4%	($1.3, $3.2)	0.00
PY 5	$34	$32	$3.2*** ($0.7)	10.3%	($2.1, $4.3)	0.00
PY 1 through 5	$30	$29	$2.1*** ($0.4)	7.6%	($1.4, $2.9)	0.00
Durable medical equipment expenditures						
Baseline	$22	$21	NA	NA	NA	NA
PY 1	$21	$19	$0.0 ($0.3)	0.0%	(-$0.4, $0.4)	0.99
PY 2	$23	$22	-$0.3 ($0.3)	-1.2%	(-$0.8, $0.2)	0.32
PY 3	$24	$24	-$0.4 ($0.3)	-1.6%	(-$0.9, $0.1)	0.21
PY 4	$26	$25	$0.0 ($0.4)	0.0%	(-$0.6, $0.6)	0.98
PY 5	$27	$26	-$0.3 ($0.4)	-1.1%	(-$0.9, $0.3)	0.44
PY 1 through 5	$24	$23	-$0.2 ($0.3)	-0.8%	(-$0.6, $0.2)	0.46

Table 5.A.2.1a. ***(continued)***

	Track 1—Overall					
	CPC+ mean[a]	C mean[a]	Impact estimate[b] (SE)	Percentage impact[c]	90 percent confidence interval	*p*-Value
Unweighted sample sizes[m]						
Number of practices	1,373	5,243				
Number of beneficiaries	1,549,585	5,347,499				
Number of beneficiary-years	5,916,394	20,150,090				

Source: Mathematica's analysis of Medicare claims data from January 2013 through December 2021.

Notes: Although this table indicates which estimates are statistically significant, when we interpret evidence, we combine evidence from the magnitude of the effect, the *p*-values, findings on related outcomes, subgroups, sensitivity tests, and other data sources about model implementation.

[a] We report the actual, unadjusted averages in the baseline period which are similar for the CPC+ and comparison groups due to matching. In the intervention periods, the comparison group mean is computed by subtracting the regression adjusted difference between the CPC+ and comparison means in each time period from the CPC+ mean in that same time period.

[b] Each impact estimate is regression-adjusted using a difference-in-differences analysis that reflects the difference of the average outcome for Medicare FFS beneficiaries attributed to CPC+ practices in the first five years of CPC+ to the average outcome in the baseline year, relative to the same difference over time for Medicare FFS beneficiaries attributed to comparison practices, while controlling for beneficiary characteristics, practice fixed effects and COVID-19 related controls.

[c] We calculated percentage impacts relative to what the CPC+ mean would have been in Program Years 1 through 5 (separately and combined) in the absence of the intervention—that is, the unadjusted CPC+ mean minus the impact estimate.

[d] Expenditures for Part A and Part B services in PY 3, PY 4, and PY 5 include QPP payment adjustments, based on practitioner performance two years before. They are applicable for both CPC+ and comparison practices. The adjustments are composed of (1) MIPS adjustments, which are applied directly to physician and outpatient claims (as a percentage of the charges on the claims); and (2) lump sum incentive payments to eligible practitioners who participated in Advanced APMs in 2017, 2018, and 2019 (calculated based on 2018, 2019, and 2020 claims for these practitioners, respectively). The first QPP adjustments were paid in PY 3 (two years after the start of QPP), so there are no QPP payments in PYs 1 and 2.

[e] We determine SSP ACO participation status based on participation at the beginning of PY 1 (January 1, 2017). However, over time, CPC+ practices may join or leave SSP, resulting in a small subset of SSP practices receiving the Performance-based Incentive Payments and a small subset of non-SSP practices receiving the shared savings payments. This is reflected in the impact estimates.

[f] The sum of expenditures by claim type does not equal the total expenditures for Part A and B services without enhanced payments in PY 3, PY 4, and PY 5 because the total expenditures include lump-sum incentive payments that are not applied at the claim level and are instead paid out directly to eligible practitioners who participated in Advanced APMs in 2017, 2018 and 2019.

[g] Acute inpatient care includes short-stay acute hospital admissions and admissions to CAHs. Expenditures for non-acute hospital admissions other than those for inpatient rehabilitation, such as psychiatric hospital admissions, are included in inpatient expenditures but not shown separately.

[h] Post-acute care expenditures include expenditures on home health, long-term care, skilled nursing facility, and inpatient rehabilitation. These are not a sub-category of inpatient expenditures.

[i] Expenditures, with QPP payment adjustments, on outpatient ED visits include professional (which is part of expenditures for physician and nonphysician Part B noninstitutional services) and facility fees, as well as payments for observation stays.

[j] Expenditures, with QPP payment adjustments, on Part B noninstitutional services include expenditures for (1) ambulatory primary care visits, (2) ambulatory specialist visits, and (3) non-ambulatory physician visits as well as services provided by other noninstitutional providers. (We only show the first two categories separately in the table).

[k] We define the assigned practice for the baseline period as the first practice to which a beneficiary was attributed during the baseline period, and the assigned practice for the intervention period as the first practice that the beneficiary was attributed to during the intervention period.

[l] Laboratory and imaging services were identified in both the carrier and outpatient claim files.

[m] After accounting for weights that adjust for matching and time observed in Medicare FFS, the effective sample sizes fall but are still substantial. For the comparison group, the effective sample size is 45 percent of the actual sample size. The effective sample size for the CPC+ group is 96 percent of the actual sample size because it is affected only by time observed (and not by the matching weights).

*/**/*** Significantly different from zero at the 0.10/0.05/0.01 level, two-tailed test.

NA = not applicable because the difference-in-differences impact estimate cannot be calculated at baseline.

ACO = Accountable Care Organization; APM = Alternative Payment Model; C = comparison; CAH = critical access hospital; ED = emergency department; FFS = fee-for-service; MIPS = Merit-based Incentive Payment System; NA = not applicable; PY = Program Year; QPP = Quality Payment Program; SE = standard error.

Table 5.A.2.1b. Regression-adjusted means and estimated impacts of CPC+ on selected Medicare expenditure outcomes for attributed Medicare FFS beneficiaries, by program year, and average across the five program years, Track 1 by SSP status

	Track 1—SSP						Track 1—Non-SSP						
	CPC+ mean[a]	C mean[a]	Impact estimate[b] (SE)	Percentage impact[c]	90 percent confidence interval	*p*-Value	CPC+ mean[a]	C mean[a]	Impact estimate[b] (SE)	Percentage impact[c]	90 percent confidence interval	*p*-Value	*p*-Value for SSP vs. non-SSP significance
Medicare expenditures (per beneficiary per month)													
Medicare Part A and B expenditures without enhanced payments for CPC+ and SSP[d]													
Baseline	$906	$905	NA	NA	NA	NA	$854	$861	NA	NA	NA	NA	NA
PY 1	$924	$921	$1.4 ($4.5)	0.2%	(-$6.1, $8.9)	0.76	$874	$871	$8.6* ($5.1)	1.0%	($0.2, $17.0)	0.09	0.29
PY 2	$975	$973	$0.0 ($4.9)	0.0%	(-$8.0, $8.0)	0.99	$923	$921	$8.2 ($5.3)	0.9%	(-$0.6, $16.9)	0.12	0.26
PY 3	$1,017	$1,024	-$8.2 ($5.5)	-0.8%	(-$17.2, $0.9)	0.14	$971	$963	$13.7** ($6.1)	1.4%	($3.6, $23.8)	0.03	0.01
PY 4	$969	$983	-$15.1** ($6.5)	-1.5%	(-$25.9, -$4.4)	0.02	$929	$925	$10.0 ($6.2)	1.1%	(-$0.3, $20.2)	0.11	0.01
PY 5	$1,073	$1,092	-$19.5*** ($6.6)	-1.8%	(-$30.3, -$8.6)	0.00	$1,009	$1,006	$9.8 ($6.7)	1.0%	(-$1.3, $20.9)	0.15	0.00
PY 1 through 5	$994	$1,000	-$7.8* ($4.5)	-0.8%	(-$15.3, -$0.4)	0.08	$944	$940	$10.1** ($4.8)	1.1%	($2.1, $18.1)	0.04	0.01
Medicare Part A and B expenditures including care management fees													
Baseline	$906	$905	NA	NA	NA	NA	$854	$861	NA	NA	NA	NA	NA
PY 1	$938	$921	$15.4*** ($4.5)	1.7%	($7.9, $22.8)	0.00	$887	$871	$22.3*** ($5.1)	2.6%	($13.9, $30.7)	0.00	0.31
PY 2	$987	$973	$12.7*** ($4.9)	1.3%	($4.7, $20.7)	0.01	$935	$921	$20.7*** ($5.3)	2.3%	($12.0, $29.4)	0.00	0.27
PY 3	$1,029	$1,024	$3.7 ($5.5)	0.4%	(-$5.3, $12.7)	0.50	$982	$963	$25.4*** ($6.1)	2.7%	($15.2, $35.5)	0.00	0.01
PY 4	$980	$983	-$4.1 ($6.5)	-0.4%	(-$14.8, $6.6)	0.53	$940	$925	$21.0*** ($6.2)	2.3%	($10.7, $31.2)	0.00	0.01
PY 5	$1,084	$1,092	-$9.1 ($6.6)	-0.8%	(-$20.0, $1.7)	0.17	$1,019	$1,006	$19.6*** ($6.8)	2.0%	($8.4, $30.7)	0.00	0.00
PY 1 through 5	$1,006	$1,000	$4.2 ($4.5)	0.4%	(-$3.3, $11.6)	0.36	$956	$940	$21.8*** ($4.9)	2.3%	($13.8, $29.8)	0.00	0.01
Medicare Part A and B expenditures including care management fees, Performance-based Incentive Payments, and shared savings payments to SSP ACOs[e]													
Baseline	$910	$908	NA	NA	NA	NA	$855	$861	NA	NA	NA	NA	NA
PY 1	$943	$926	$15.7*** ($4.5)	1.7%	($8.2, $23.2)	0.00	$889	$871	$23.7*** ($5.1)	2.7%	($15.3, $32.1)	0.00	0.24
PY 2	$994	$978	$13.8*** ($4.8)	1.4%	($5.9, $21.8)	0.00	$938	$922	$22.0*** ($5.3)	2.4%	($13.3, $30.7)	0.00	0.25
PY 3	$1,037	$1,031	$4.7 ($5.4)	0.5%	(-$4.1, $13.5)	0.38	$985	$966	$25.5*** ($6.1)	2.7%	($15.4, $35.6)	0.00	0.01
PY 4	$988	$994	-$7.6 ($6.5)	-0.8%	(-$18.3, $3.0)	0.24	$944	$929	$20.7*** ($6.2)	2.2%	($10.5, $31.0)	0.00	0.00

Table 5.A.2.1b. ***(continued)***

	Track 1—SSP						Track 1—Non-SSP						
	CPC+ mean[a]	C mean[a]	Impact estimate[b] (SE)	Percentage impact[c]	90 percent confidence interval	*p*-Value	CPC+ mean[a]	C mean[a]	Impact estimate[b] (SE)	Percentage impact[c]	90 percent confidence interval	*p*-Value	*p*-Value for SSP vs. non-SSP significance
PY 5	$1,090	$1,101	-$12.2* ($6.5)	-1.1%	(-$23.0, -$1.5)	0.06	$1,023	$1,010	$18.9*** ($6.8)	1.9%	($7.8, $30.1)	0.01	0.00
PY 1 through 5	$1,013	$1,008	$3.5 ($4.4)	0.3%	(-$3.8, $10.8)	0.43	$958	$942	$22.2*** ($4.9)	2.4%	($14.2, $30.2)	0.00	0.00
Medicare expenditures by service category (per beneficiary per month)[f]													
Inpatient expenditures													
Baseline	$318	$322	NA	NA	NA	NA	$303	$314	NA	NA	NA	NA	NA
PY 1	$323	$326	$0.6 ($3.0)	0.2%	(-$4.3, $5.5)	0.84	$308	$313	$5.2 ($3.5)	1.7%	(-$0.6, $11.0)	0.14	0.32
PY 2	$331	$335	-$0.5 ($3.1)	-0.1%	(-$5.6, $4.6)	0.88	$312	$321	$2.2 ($3.4)	0.7%	(-$3.4, $7.7)	0.52	0.57
PY 3	$340	$351	-$6.5* ($3.5)	-1.9%	(-$12.2, -$0.8)	0.06	$324	$332	$2.2 ($3.7)	0.7%	(-$3.9, $8.3)	0.55	0.09
PY 4	$322	$338	-$12.3*** ($3.7)	-3.7%	(-$18.3, -$6.2)	0.00	$309	$317	$2.3 ($3.7)	0.7%	(-$3.9, $8.4)	0.54	0.01
PY 5	$341	$357	-$11.8*** ($3.9)	-3.3%	(-$18.2, -$5.5)	0.00	$319	$326	$3.6 ($3.9)	1.1%	(-$2.8, $9.9)	0.35	0.00
PY 1 through 5	$332	$341	-$5.8** ($2.7)	-1.7%	(-$10.3, -$1.4)	0.03	$315	$322	$3.0 ($3.0)	1.0%	(-$1.9, $8.0)	0.31	0.03
Expenditures for acute inpatient care[g]													
Baseline	$282	$285	NA	NA	NA	NA	$268	$278	NA	NA	NA	NA	NA
PY 1	$285	$290	-$1.4 ($2.6)	-0.5%	(-$5.7, $2.9)	0.59	$273	$279	$4.0 ($3.1)	1.5%	(-$1.1, $9.2)	0.20	0.18
PY 2	$292	$298	-$2.6 ($2.7)	-0.9%	(-$7.0, $1.9)	0.34	$276	$287	-$0.4 ($3.0)	-0.2%	(-$5.4, $4.5)	0.89	0.59
PY 3	$302	$314	-$8.1*** ($3.0)	-2.6%	(-$13.1, -$3.2)	0.01	$287	$297	-$0.1 ($3.2)	0.0%	(-$5.4, $5.2)	0.97	0.07
PY 4	$286	$302	-$12.3*** ($3.2)	-4.1%	(-$17.6, -$6.9)	0.00	$273	$284	-$1.1 ($3.3)	-0.4%	(-$6.5, $4.4)	0.75	0.02
PY 5	$306	$320	-$10.5*** ($3.5)	-3.3%	(-$16.2, -$4.8)	0.00	$283	$291	$1.3 ($3.5)	0.5%	(-$4.5, $7.0)	0.71	0.02
PY 1 through 5	$295	$305	-$6.8*** ($2.4)	-2.3%	(-$10.7, -$2.9)	0.00	$278	$288	$0.7 ($2.7)	0.3%	(-$3.7, $5.1)	0.79	0.03
Expenditures for acute surgical hospitalizations													
Baseline	$152	$151	NA	NA	NA	NA	$143	$146	NA	NA	NA	NA	NA
PY 1	$151	$150	$0.4 ($2.1)	0.3%	(-$3.0, $3.8)	0.85	$145	$143	$5.0** ($2.3)	3.6%	($1.2, $8.8)	0.03	0.13
PY 2	$153	$153	-$0.3 ($2.0)	-0.2%	(-$3.6, $3.0)	0.88	$144	$145	$1.2 ($2.1)	0.8%	(-$2.3, $4.7)	0.58	0.62
PY 3	$158	$159	-$2.4 ($2.2)	-1.5%	(-$6.0, $1.2)	0.28	$150	$151	$1.5 ($2.2)	1.0%	(-$2.2, $5.2)	0.49	0.21

Table 5.A.2.1b. ***(continued)***

	Track 1—SSP						Track 1—Non-SSP						
	CPC+ mean[a]	C mean[a]	Impact estimate[b] (SE)	Percentage impact[c]	90 percent confidence interval	*p*-Value	CPC+ mean[a]	C mean[a]	Impact estimate[b] (SE)	Percentage impact[c]	90 percent confidence interval	*p*-Value	*p*-Value for SSP vs. non-SSP significance
PY 4	$141	$144	-$4.0* ($2.2)	-2.8%	(-$7.6, -$0.4)	0.07	$136	$136	$2.5 ($2.2)	1.9%	(-$1.1, $6.2)	0.25	0.03
PY 5	$147	$151	-$4.8** ($2.4)	-3.2%	(-$8.7, -$0.9)	0.04	$138	$139	$2.0 ($2.4)	1.5%	(-$1.9, $5.9)	0.40	0.04
PY 1 through 5	$150	$151	-$2.1 ($1.7)	-1.4%	(-$4.9, $0.7)	0.21	$142	$143	$2.4 ($1.8)	1.7%	(-$0.6, $5.4)	0.19	0.07
Expenditures for acute medical hospitalizations													
Baseline	$130	$134	NA	NA	NA	NA	$125	$132	NA	NA	NA	NA	NA
PY 1	$134	$140	-$1.8 ($1.3)	-1.3%	(-$3.9, $0.3)	0.17	$128	$136	-$1.0 ($1.6)	-0.7%	(-$3.5, $1.6)	0.54	0.68
PY 2	$139	$146	-$2.3 ($1.5)	-1.6%	(-$4.7, $0.2)	0.13	$132	$141	-$1.6 ($1.6)	-1.2%	(-$4.3, $1.1)	0.33	0.76
PY 3	$144	$155	-$5.8*** ($1.7)	-3.8%	(-$8.5, -$3.0)	0.00	$137	$146	-$1.7 ($1.8)	-1.2%	(-$4.6, $1.3)	0.36	0.10
PY 4	$146	$158	-$8.3*** ($1.9)	-5.4%	(-$11.3, -$5.2)	0.00	$137	$148	-$3.6* ($2.1)	-2.6%	(-$7.0, -$0.2)	0.08	0.10
PY 5	$158	$168	-$5.7*** ($2.1)	-3.5%	(-$9.1, -$2.3)	0.01	$144	$152	-$0.7 ($2.1)	-0.5%	(-$4.1, $2.7)	0.73	0.09
PY 1 through 5	$145	$154	-$4.7*** ($1.3)	-3.1%	(-$6.9, -$2.5)	0.00	$136	$145	-$1.7 ($1.5)	-1.2%	(-$4.1, $0.7)	0.25	0.13
Inpatient rehabilitation facility expenditures													
Baseline	$21	$21	NA	NA	NA	NA	$20	$21	NA	NA	NA	NA	NA
PY 1	$22	$21	$0.6 ($0.6)	2.7%	(-$0.4, $1.6)	0.33	$21	$21	$1.0* ($0.6)	5.2%	($0.0, $2.1)	0.10	0.61
PY 2	$23	$22	$1.0 ($0.7)	4.6%	(-$0.1, $2.2)	0.13	$22	$21	$2.1*** ($0.7)	10.3%	($1.0, $3.2)	0.00	0.27
PY 3	$24	$23	$0.3 ($0.7)	1.2%	(-$0.9, $1.5)	0.70	$23	$22	$2.0*** ($0.7)	9.5%	($0.8, $3.2)	0.01	0.09
PY 4	$23	$23	-$0.4 ($0.7)	-1.6%	(-$1.6, $0.8)	0.61	$24	$22	$3.0*** ($0.7)	14.5%	($1.8, $4.1)	0.00	0.00
PY 5	$24	$25	-$0.7 ($0.8)	-2.7%	(-$1.9, $0.6)	0.38	$25	$24	$2.1*** ($0.8)	9.4%	($0.9, $3.4)	0.01	0.01
PY 1 through 5	$23	$23	$0.2 ($0.6)	0.9%	(-$0.7, $1.1)	0.71	$23	$22	$2.1*** ($0.5)	9.7%	($1.1, $3.0)	0.00	0.02
Post-acute care expenditures[h]													
Baseline	$115	$118	NA	NA	NA	NA	$105	$107	NA	NA	NA	NA	NA
PY 1	$113	$115	$1.1 ($1.6)	0.9%	(-$1.5, $3.6)	0.50	$104	$104	$1.9 ($1.7)	1.9%	(-$0.9, $4.7)	0.26	0.71
PY 2	$112	$114	$0.9 ($1.6)	0.8%	(-$1.7, $3.6)	0.56	$102	$103	$1.1 ($1.6)	1.1%	(-$1.6, $3.7)	0.51	0.95
PY 3	$109	$115	-$2.7 ($1.7)	-2.4%	(-$5.5, $0.0)	0.10	$103	$102	$2.3 ($1.9)	2.2%	(-$0.8, $5.4)	0.23	0.05

Table 5.A.2.1b. ***(continued)***

	Track 1—SSP						Track 1—Non-SSP						
	CPC+ mean[a]	C mean[a]	Impact estimate[b] (SE)	Percentage impact[c]	90 percent confidence interval	*p*-Value	CPC+ mean[a]	C mean[a]	Impact estimate[b] (SE)	Percentage impact[c]	90 percent confidence interval	*p*-Value	*p*-Value for SSP vs. non-SSP significance
PY 4	$98	$102	-$1.4 ($2.0)	-1.4%	(-$4.7, $1.8)	0.47	$92	$92	$2.5 ($1.8)	2.8%	(-$0.5, $5.6)	0.17	0.14
PY 5	$96	$101	-$2.8 ($1.8)	-2.9%	(-$5.8, $0.2)	0.12	$88	$88	$1.7 ($1.9)	1.9%	(-$1.5, $4.8)	0.38	0.09
PY 1 through 5	$105	$109	-$1.0 ($1.4)	-0.9%	(-$3.3, $1.3)	0.49	$98	$97	$1.9 ($1.4)	2.0%	(-$0.5, $4.3)	0.19	0.16
Acute inpatient and post-acute care expenditures combined[h]													
Baseline	$397	$403	NA	NA	NA	NA	$373	$385	NA	NA	NA	NA	NA
PY 1	$398	$405	-$0.4 ($3.6)	-0.1%	(-$6.2, $5.5)	0.92	$377	$382	$6.0 ($4.1)	1.6%	(-$0.8, $12.7)	0.15	0.24
PY 2	$405	$413	-$1.6 ($3.7)	-0.4%	(-$7.7, $4.4)	0.65	$379	$389	$0.6 ($3.9)	0.2%	(-$5.8, $7.1)	0.87	0.67
PY 3	$411	$429	-$10.9*** ($4.1)	-2.6%	(-$17.6, -$4.1)	0.01	$390	$399	$2.1 ($4.3)	0.6%	(-$5.0, $9.3)	0.62	0.03
PY 4	$384	$404	-$13.7*** ($4.6)	-3.4%	(-$21.2, -$6.2)	0.00	$365	$375	$1.5 ($4.4)	0.4%	(-$5.7, $8.6)	0.74	0.02
PY 5	$401	$421	-$13.3*** ($4.6)	-3.2%	(-$20.8, -$5.8)	0.00	$371	$379	$3.0 ($4.6)	0.8%	(-$4.7, $10.6)	0.52	0.01
PY 1 through 5	$400	$414	-$7.8** ($3.3)	-1.9%	(-$13.2, -$2.4)	0.02	$376	$385	$2.6 ($3.5)	0.7%	(-$3.1, $8.3)	0.45	0.03
Outpatient expenditures													
Baseline	$164	$168	NA	NA	NA	NA	$167	$171	NA	NA	NA	NA	NA
PY 1	$176	$179	$0.8 ($1.1)	0.4%	(-$1.1, $2.6)	0.50	$177	$181	$0.9 ($1.3)	0.5%	(-$1.3, $3.0)	0.51	0.95
PY 2	$197	$200	$0.9 ($1.4)	0.5%	(-$1.4, $3.2)	0.53	$201	$203	$2.5 ($1.8)	1.3%	(-$0.4, $5.4)	0.16	0.48
PY 3	$211	$216	-$1.1 ($1.6)	-0.5%	(-$3.8, $1.6)	0.50	$216	$218	$3.5* ($2.1)	1.7%	($0.1, $7.0)	0.09	0.08
PY 4	$201	$205	-$1.1 ($2.0)	-0.5%	(-$4.3, $2.1)	0.57	$206	$211	$0.2 ($2.4)	0.1%	(-$3.8, $4.3)	0.92	0.67
PY 5	$231	$240	-$4.6** ($2.3)	-2.0%	(-$8.4, -$0.9)	0.04	$233	$242	-$3.8 ($2.6)	-1.6%	(-$8.1, $0.5)	0.14	0.82
PY 1 through 5	$204	$209	-$0.9 ($1.3)	-0.5%	(-$3.1, $1.2)	0.48	$208	$212	$0.7 ($1.7)	0.3%	(-$2.1, $3.5)	0.68	0.45
Expenditures for outpatient ED visits, including observation stays[i]													
Baseline	$25	$26	NA	NA	NA	NA	$25	$26	NA	NA	NA	NA	NA
PY 1	$26	$27	$0.2 ($0.2)	0.8%	(-$0.2, $0.6)	0.39	$27	$28	$0.0 ($0.3)	-0.1%	(-$0.5, $0.4)	0.91	0.52
PY 2	$28	$28	$0.4 ($0.3)	1.6%	(-$0.1, $0.9)	0.16	$28	$30	-$0.4 ($0.4)	-1.4%	(-$1.0, $0.2)	0.28	0.08
PY 3	$28	$30	-$0.1 ($0.3)	-0.3%	(-$0.6, $0.4)	0.75	$30	$31	-$0.1 ($0.4)	-0.3%	(-$0.7, $0.6)	0.84	0.97

Table 5.A.2.1b. ***(continued)***

	Track 1—SSP						Track 1—Non-SSP						
	CPC+ mean[a]	C mean[a]	Impact estimate[b] (SE)	Percentage impact[c]	90 percent confidence interval	*p*-Value	CPC+ mean[a]	C mean[a]	Impact estimate[b] (SE)	Percentage impact[c]	90 percent confidence interval	*p*-Value	*p*-Value for SSP vs. non-SSP significance
PY 4	$23	$24	$0.0 ($0.3)	-0.1%	(-$0.6, $0.5)	0.93	$25	$26	$0.0 ($0.4)	-0.1%	(-$0.7, $0.6)	0.96	0.98
PY 5	$27	$29	-$0.6 ($0.4)	-2.2%	(-$1.2, $0.0)	0.11	$29	$31	-$0.8 ($0.5)	-2.8%	(-$1.7, $0.1)	0.12	0.72
PY 1 through 5	$27	$28	$0.0 ($0.2)	0.0%	(-$0.4, $0.4)	1.00	$28	$29	-$0.3 ($0.3)	-1.0%	(-$0.8, $0.3)	0.40	0.50
Expenditures for physician and nonphysician Part B noninstitutional services in any setting[j]													
Baseline	$269	$254	NA	NA	NA	NA	$238	$229	NA	NA	NA	NA	NA
PY 1	$272	$259	-$1.3 ($1.1)	-0.5%	(-$3.2, $0.5)	0.22	$244	$233	$1.4 ($1.1)	0.6%	(-$0.5, $3.3)	0.22	0.08
PY 2	$289	$275	$0.1 ($1.3)	0.0%	(-$2.1, $2.3)	0.92	$259	$247	$2.6* ($1.6)	1.0%	($0.1, $5.2)	0.09	0.22
PY 3	$305	$289	$1.4 ($1.6)	0.5%	(-$1.2, $4.0)	0.38	$274	$260	$4.2** ($1.9)	1.6%	($1.1, $7.3)	0.03	0.25
PY 4	$284	$269	$1.4 ($1.8)	0.5%	(-$1.5, $4.3)	0.44	$258	$243	$5.5** ($2.2)	2.2%	($2.0, $9.1)	0.01	0.14
PY 5	$333	$319	-$0.1 ($2.0)	0.0%	(-$3.4, $3.2)	0.98	$297	$282	$5.9** ($2.5)	2.0%	($1.8, $10.0)	0.02	0.06
PY 1 through 5	$297	$283	$0.3 ($1.3)	0.1%	(-$1.8, $2.4)	0.80	$267	$254	$3.9** ($1.6)	1.5%	($1.3, $6.6)	0.01	0.08
Expenditures for ambulatory visits with primary care practitioners													
Baseline	$24	$25	NA	NA	NA	NA	$23	$23	NA	NA	NA	NA	NA
PY 1	$24	$25	-$0.2 ($0.1)	-0.7%	(-$0.4, $0.0)	0.20	$24	$24	-$0.2 ($0.2)	-0.7%	(-$0.4, $0.1)	0.26	0.97
PY 2	$25	$26	-$0.1 ($0.2)	-0.5%	(-$0.4, $0.1)	0.42	$25	$25	$0.1 ($0.2)	0.3%	(-$0.3, $0.4)	0.73	0.43
PY 3	$27	$28	-$0.1 ($0.2)	-0.5%	(-$0.5, $0.2)	0.43	$26	$26	-$0.2 ($0.2)	-0.7%	(-$0.6, $0.2)	0.46	0.91
PY 4	$24	$25	-$0.1 ($0.2)	-0.6%	(-$0.5, $0.2)	0.54	$24	$24	-$0.5 ($0.3)	-2.0%	(-$1.0, $0.0)	0.11	0.37
PY 5	$31	$32	-$0.5 ($0.3)	-1.7%	(-$1.1, $0.0)	0.10	$30	$31	-$0.6* ($0.4)	-2.1%	(-$1.2, -$0.1)	0.07	0.81
PY 1 through 5	$26	$27	-$0.2 ($0.2)	-0.8%	(-$0.5, $0.1)	0.19	$26	$26	-$0.3 ($0.2)	-1.1%	(-$0.6, $0.1)	0.21	0.83
Expenditures for ambulatory visits with primary care practitioners at assigned practice[k]													
Baseline	$17	$17	NA	NA	NA	NA	$17	$16	NA	NA	NA	NA	NA
PY 1	$17	$17	-$0.1 ($0.1)	-0.8%	(-$0.4, $0.1)	0.36	$17	$16	$0.0 ($0.2)	-0.1%	(-$0.3, $0.2)	0.90	0.60
PY 2	$15	$16	$0.0 ($0.2)	0.2%	(-$0.3, $0.4)	0.88	$15	$14	$0.3 ($0.2)	1.9%	(-$0.1, $0.7)	0.24	0.44
PY 3	$15	$15	$0.2 ($0.2)	1.1%	(-$0.2, $0.6)	0.49	$15	$14	$0.0 ($0.3)	0.2%	(-$0.5, $0.5)	0.91	0.74

Table 5.A.2.1b. ***(continued)***

	Track 1—SSP						Track 1—Non-SSP						
	CPC+ mean[a]	C mean[a]	Impact estimate[b] (SE)	Percentage impact[c]	90 percent confidence interval	*p*-Value	CPC+ mean[a]	C mean[a]	Impact estimate[b] (SE)	Percentage impact[c]	90 percent confidence interval	*p*-Value	*p*-Value for SSP vs. non-SSP significance
PY 4	$12	$12	-$0.1 ($0.3)	-0.5%	(-$0.5, $0.4)	0.83	$12	$11	-$0.1 ($0.4)	-0.5%	(-$0.7, $0.6)	0.87	1.00
PY 5	$14	$15	-$0.2 ($0.4)	-1.1%	(-$0.8, $0.5)	0.70	$15	$14	$0.1 ($0.4)	0.8%	(-$0.6, $0.8)	0.79	0.65
PY 1 through 5	$15	$15	$0.0 ($0.2)	-0.2%	(-$0.4, $0.3)	0.91	$15	$14	$0.1 ($0.3)	0.5%	(-$0.4, $0.5)	0.78	0.77
Expenditures for ambulatory visits with primary care practitioners at non-assigned practice[k]													
Baseline	$7	$7	NA	NA	NA	NA	$7	$7	NA	NA	NA	NA	NA
PY 1	$8	$8	$0.0 ($0.1)	-0.5%	(-$0.2, $0.1)	0.65	$8	$8	-$0.1 ($0.1)	-1.9%	(-$0.3, $0.0)	0.13	0.42
PY 2	$10	$11	-$0.2 ($0.2)	-1.7%	(-$0.5, $0.1)	0.32	$10	$11	-$0.2 ($0.2)	-2.0%	(-$0.5, $0.1)	0.19	0.87
PY 3	$11	$12	-$0.3 ($0.2)	-2.6%	(-$0.6, $0.0)	0.11	$11	$12	-$0.2 ($0.2)	-1.9%	(-$0.5, $0.1)	0.26	0.74
PY 4	$12	$13	-$0.1 ($0.2)	-0.7%	(-$0.5, $0.3)	0.74	$12	$13	-$0.4 ($0.3)	-3.5%	(-$0.9, $0.0)	0.14	0.36
PY 5	$16	$17	-$0.4 ($0.3)	-2.2%	(-$0.9, $0.1)	0.24	$16	$17	-$0.8** ($0.3)	-4.6%	(-$1.3, -$0.2)	0.02	0.40
PY 1 through 5	$12	$12	-$0.2 ($0.2)	-1.6%	(-$0.5, $0.1)	0.25	$11	$13	-$0.4** ($0.2)	-3.0%	(-$0.6, -$0.1)	0.05	0.53
Expenditures for ambulatory visits with specialists													
Baseline	$28	$26	NA	NA	NA	NA	$23	$22	NA	NA	NA	NA	NA
PY 1	$27	$25	$0.0 ($0.1)	-0.2%	(-$0.2, $0.1)	0.58	$23	$22	$0.2*** ($0.1)	1.1%	($0.1, $0.4)	0.01	0.02
PY 2	$28	$26	$0.1 ($0.1)	0.4%	(-$0.1, $0.3)	0.32	$23	$22	$0.2** ($0.1)	1.1%	($0.1, $0.4)	0.03	0.41
PY 3	$28	$26	$0.0 ($0.1)	0.1%	(-$0.2, $0.2)	0.90	$23	$23	$0.2 ($0.1)	0.7%	(-$0.1, $0.4)	0.24	0.46
PY 4	$24	$23	-$0.2 ($0.1)	-0.9%	(-$0.4, $0.0)	0.14	$20	$19	$0.3* ($0.1)	1.4%	($0.0, $0.5)	0.06	0.02
PY 5	$33	$32	-$0.6*** ($0.2)	-1.8%	(-$0.9, -$0.3)	0.00	$26	$26	$0.2 ($0.2)	0.7%	(-$0.1, $0.5)	0.29	0.00
PY 1 through 5	$28	$26	-$0.1 ($0.1)	-0.5%	(-$0.3, $0.0)	0.23	$23	$23	$0.2* ($0.1)	1.0%	($0.0, $0.4)	0.05	0.03
Expenditures on laboratory services[l]													
Baseline	$26	$25	NA	NA	NA	NA	$23	$24	NA	NA	NA	NA	NA
PY 1	$26	$25	-$0.2 ($0.1)	-0.8%	(-$0.4, $0.0)	0.16	$24	$24	$0.2 ($0.2)	0.9%	($0.0, $0.5)	0.14	0.04
PY 2	$28	$27	$0.1 ($0.2)	0.3%	(-$0.2, $0.4)	0.67	$26	$26	$0.4* ($0.2)	1.7%	($0.1, $0.8)	0.05	0.20
PY 3	$29	$28	-$0.2 ($0.3)	-0.9%	(-$0.7, $0.2)	0.35	$26	$27	$0.4 ($0.2)	1.4%	(-$0.1, $0.8)	0.15	0.10

Table 5.A.2.1b. ***(continued)***

	Track 1—SSP						Track 1—Non-SSP						
	CPC+ mean[a]	C mean[a]	Impact estimate[b] (SE)	Percentage impact[c]	90 percent confidence interval	*p*-Value	CPC+ mean[a]	C mean[a]	Impact estimate[b] (SE)	Percentage impact[c]	90 percent confidence interval	*p*-Value	*p*-Value for SSP vs. non-SSP significance
PY 4	$30	$30	-$0.6** ($0.3)	-2.0%	(-$1.1, -$0.1)	0.05	$28	$28	$0.2 ($0.3)	0.6%	(-$0.3, $0.6)	0.55	0.06
PY 5	$36	$36	-$1.6*** ($0.5)	-4.2%	(-$2.5, -$0.7)	0.00	$32	$33	$0.1 ($0.3)	0.3%	(-$0.5, $0.6)	0.79	0.01
PY 1 through 5	$30	$29	-$0.5** ($0.2)	-1.5%	(-$0.8, -$0.1)	0.02	$27	$28	$0.3 ($0.2)	0.9%	(-$0.1, $0.6)	0.20	0.01
Expenditures on imaging services[l]													
Baseline	$49	$47	NA	NA	NA	NA	$44	$44	NA	NA	NA	NA	NA
PY 1	$50	$49	-$0.9** ($0.3)	-1.7%	(-$1.4, -$0.3)	0.01	$45	$45	-$0.1 ($0.3)	-0.3%	(-$0.7, $0.4)	0.71	0.13
PY 2	$53	$52	-$0.5 ($0.4)	-0.9%	(-$1.1, $0.1)	0.18	$48	$48	-$0.3 ($0.4)	-0.5%	(-$0.9, $0.4)	0.51	0.68
PY 3	$56	$55	-$0.7* ($0.4)	-1.3%	(-$1.4, -$0.1)	0.07	$51	$51	$0.4 ($0.4)	0.8%	(-$0.3, $1.1)	0.34	0.05
PY 4	$50	$49	-$0.5 ($0.4)	-0.9%	(-$1.2, $0.2)	0.28	$46	$46	$0.4 ($0.4)	0.9%	(-$0.3, $1.1)	0.34	0.15
PY 5	$56	$56	-$1.1** ($0.5)	-1.9%	(-$1.8, -$0.3)	0.02	$52	$52	-$0.4 ($0.5)	-0.8%	(-$1.2, $0.4)	0.39	0.30
PY 1 through 5	$53	$52	-$0.7** ($0.3)	-1.3%	(-$1.2, -$0.2)	0.02	$49	$49	$0.0 ($0.3)	0.0%	(-$0.5, $0.5)	0.99	0.11
Skilled nursing facility expenditures													
Baseline	$71	$72	NA	NA	NA	NA	$63	$64	NA	NA	NA	NA	NA
PY 1	$69	$70	$0.3 ($1.0)	0.4%	(-$1.3, $1.9)	0.77	$61	$61	$0.6 ($1.0)	0.9%	(-$1.1, $2.2)	0.58	0.84
PY 2	$68	$70	-$0.5 ($1.0)	-0.7%	(-$2.1, $1.2)	0.64	$61	$61	$0.2 ($1.0)	0.4%	(-$1.5, $2.0)	0.83	0.63
PY 3	$66	$70	-$2.4** ($1.1)	-3.5%	(-$4.2, -$0.6)	0.03	$60	$60	$1.7 ($1.1)	2.9%	(-$0.1, $3.5)	0.12	0.01
PY 4	$67	$72	-$4.0** ($1.9)	-5.6%	(-$7.1, -$0.8)	0.04	$60	$60	$0.9 ($1.4)	1.6%	(-$1.3, $3.2)	0.49	0.04
PY 5	$68	$73	-$3.9** ($1.6)	-5.3%	(-$6.5, -$1.2)	0.02	$61	$59	$2.7* ($1.4)	4.6%	($0.4, $5.0)	0.06	0.00
PY 1 through 5	$68	$71	-$2.0** ($1.0)	-2.9%	(-$3.7, -$0.3)	0.05	$61	$60	$1.2 ($0.9)	2.1%	(-$0.3, $2.7)	0.18	0.02
Home health expenditures													
Baseline	$40	$44	NA	NA	NA	NA	$39	$38	NA	NA	NA	NA	NA
PY 1	$40	$44	$0.0 ($0.4)	-0.1%	(-$0.7, $0.6)	0.90	$39	$38	-$0.5 ($0.4)	-1.2%	(-$1.2, $0.2)	0.24	0.44
PY 2	$40	$45	-$1.3*** ($0.4)	-3.1%	(-$2.0, -$0.6)	0.00	$39	$39	-$0.7 ($0.5)	-1.9%	(-$1.5, $0.0)	0.11	0.39
PY 3	$39	$45	-$1.7*** ($0.5)	-4.1%	(-$2.4, -$0.9)	0.00	$39	$40	-$1.4*** ($0.6)	-3.6%	(-$2.3, -$0.5)	0.01	0.75

Table 5.A.2.1b. ***(continued)***

	Track 1—SSP						Track 1—Non-SSP						
	CPC+ mean[a]	C mean[a]	Impact estimate[b] (SE)	Percentage impact[c]	90 percent confidence interval	*p*-Value	CPC+ mean[a]	C mean[a]	Impact estimate[b] (SE)	Percentage impact[c]	90 percent confidence interval	*p*-Value	*p*-Value for SSP vs. non-SSP significance
PY 4	$36	$42	-$1.7*** ($0.5)	-4.5%	(-$2.6, -$0.8)	0.00	$36	$37	-$1.9*** ($0.6)	-4.9%	(-$2.9, -$0.8)	0.00	0.82
PY 5	$38	$45	-$2.5*** ($0.6)	-6.1%	(-$3.5, -$1.5)	0.00	$38	$39	-$1.8*** ($0.6)	-4.7%	(-$2.9, -$0.8)	0.00	0.47
PY 1 through 5	$39	$44	-$1.4*** ($0.4)	-3.5%	(-$2.1, -$0.8)	0.00	$38	$39	-$1.3*** ($0.5)	-3.2%	(-$2.0, -$0.5)	0.01	0.82
Hospice expenditures													
Baseline	$22	$25	NA	NA	NA	NA	$23	$23	NA	NA	NA	NA	NA
PY 1	$24	$25	$1.4*** ($0.5)	6.2%	($0.5, $2.3)	0.01	$24	$24	$0.8 ($0.6)	3.4%	(-$0.2, $1.7)	0.19	0.43
PY 2	$27	$27	$2.1*** ($0.7)	8.3%	($1.0, $3.2)	0.00	$27	$27	$1.1 ($0.7)	4.3%	(-$0.1, $2.3)	0.13	0.32
PY 3	$31	$30	$2.9*** ($0.7)	10.4%	($1.7, $4.1)	0.00	$31	$30	$2.0** ($0.8)	6.8%	($0.6, $3.3)	0.02	0.42
PY 4	$33	$32	$3.1*** ($0.8)	10.4%	($1.7, $4.4)	0.00	$32	$31	$1.5 ($0.9)	4.7%	($0.0, $2.9)	0.10	0.18
PY 5	$34	$32	$4.1*** ($0.9)	13.5%	($2.6, $5.5)	0.00	$34	$32	$2.1** ($1.0)	6.7%	($0.5, $3.8)	0.03	0.15
PY 1 through 5	$30	$29	$2.7*** ($0.6)	9.8%	($1.7, $3.7)	0.00	$30	$29	$1.5** ($0.7)	5.3%	($0.4, $2.6)	0.02	0.19
Durable medical equipment expenditures													
Baseline	$22	$20	NA	NA	NA	NA	$22	$21	NA	NA	NA	NA	NA
PY 1	$20	$19	-$0.3 ($0.3)	-1.2%	(-$0.8, $0.3)	0.44	$21	$20	$0.3 ($0.4)	1.3%	(-$0.4, $0.9)	0.48	0.30
PY 2	$22	$22	-$0.8** ($0.4)	-3.5%	(-$1.4, -$0.2)	0.03	$24	$23	$0.3 ($0.4)	1.2%	(-$0.4, $1.0)	0.50	0.05
PY 3	$24	$23	-$0.8* ($0.4)	-3.1%	(-$1.5, -$0.1)	0.06	$25	$24	$0.0 ($0.5)	0.1%	(-$0.7, $0.8)	0.97	0.21
PY 4	$25	$24	-$0.3 ($0.5)	-1.4%	(-$1.1, $0.4)	0.46	$26	$25	$0.3 ($0.5)	1.3%	(-$0.5, $1.2)	0.52	0.33
PY 5	$26	$25	-$0.7 ($0.5)	-2.5%	(-$1.5, $0.2)	0.19	$27	$26	$0.2 ($0.5)	0.7%	(-$0.7, $1.0)	0.74	0.25
PY 1 through 5	$24	$23	-$0.6 ($0.4)	-2.4%	(-$1.2, $0.0)	0.10	$25	$24	$0.2 ($0.4)	0.9%	(-$0.4, $0.9)	0.60	0.14
Unweighted sample sizes[m]													
Number of practices	738	2,979					635	2,264					
Number of beneficiaries	798,817	3,129,830					753,337	2,233,041					
Number of beneficiary-years	3,017,546	11,762,356					2,898,848	8,387,734					

Source: Mathematica's analysis of Medicare claims data from January 2013 through December 2021.

Table 5.A.2.1b. ***(continued)***

Notes: Although this table indicates which estimates are statistically significant, when we interpret evidence, we combine evidence from the magnitude of the effect, the *p*-values, findings on related outcomes, subgroups, sensitivity tests, and other data sources about model implementation.

[a] We report the actual, unadjusted averages in the baseline period which are similar for the CPC+ and comparison groups due to matching. In the intervention periods, the comparison group mean is computed by subtracting the regression adjusted difference between the CPC+ and comparison means in each time period from the CPC+ mean in that same time period.

[b] Each impact estimate is regression-adjusted using a difference-in-differences analysis that reflects the difference of the average outcome for Medicare FFS beneficiaries attributed to CPC+ practices in the first five years of CPC+ to the average outcome in the baseline year, relative to the same difference over time for Medicare FFS beneficiaries attributed to comparison practices, while controlling for beneficiary characteristics, practice fixed effects and COVID-19 related controls.

[c] We calculated percentage impacts relative to what the CPC+ mean would have been in Program Years 1 through 5 (separately and combined) in the absence of the intervention—that is, the unadjusted CPC+ mean minus the impact estimate.

[d] Expenditures for Part A and Part B services in PY 3, PY 4, and PY 5 include QPP payment adjustments, based on practitioner performance two years before. They are applicable for both CPC+ and comparison practices. The adjustments are composed of (1) MIPS adjustments, which are applied directly to physician and outpatient claims (as a percentage of the charges on the claims); and (2) lump sum incentive payments to eligible practitioners who participated in Advanced APMs in 2017, 2018, and 2019 (calculated based on 2018, 2019, and 2020 claims for these practitioners, respectively). The first QPP adjustments were paid in PY 3 (two years after the start of QPP), so there are no QPP payments in PYs 1 and 2.

[e] We determine SSP ACO participation status based on participation at the beginning of PY 1 (January 1, 2017). However, over time, CPC+ practices may join or leave SSP, resulting in a small subset of SSP practices receiving the Performance-based Incentive Payments and a small subset of non-SSP practices receiving the shared savings payments. This is reflected in the impact estimates.

[f] The sum of expenditures by claim type does not equal the total expenditures for Part A and B services without enhanced payments in PY 3, PY 4, and PY 5 because the total expenditures include lump-sum incentive payments that are not applied at the claim level and are instead paid out directly to eligible practitioners who participated in Advanced APMs in 2017, 2018 and 2019.

[g] Acute inpatient care includes short-stay acute hospital admissions and admissions to CAHs. Expenditures for non-acute hospital admissions other than those for inpatient rehabilitation, such as psychiatric hospital admissions, are included in inpatient expenditures but not shown separately.

[h] Post-acute care expenditures include expenditures on home health, long-term care, skilled nursing facility, and inpatient rehabilitation. These are not a sub-category of inpatient expenditures.

[i] Expenditures, with QPP payment adjustments, on outpatient ED visits include professional (which is part of expenditures for physician and nonphysician Part B noninstitutional services) and facility fees, as well as payments for observation stays.

[j] Expenditures, with QPP payment adjustments, on Part B noninstitutional services include expenditures for (1) ambulatory primary care visits, (2) ambulatory specialist visits, and (3) non-ambulatory physician visits as well as services provided by other noninstitutional providers. (We only show the first two categories separately in the table).

[k] We define the assigned practice for the baseline period as the first practice to which a beneficiary was attributed during the baseline period, and the assigned practice for the intervention period as the first practice that the beneficiary was attributed to during the intervention period.

[l] Laboratory and imaging services were identified in both the carrier and outpatient claim files.

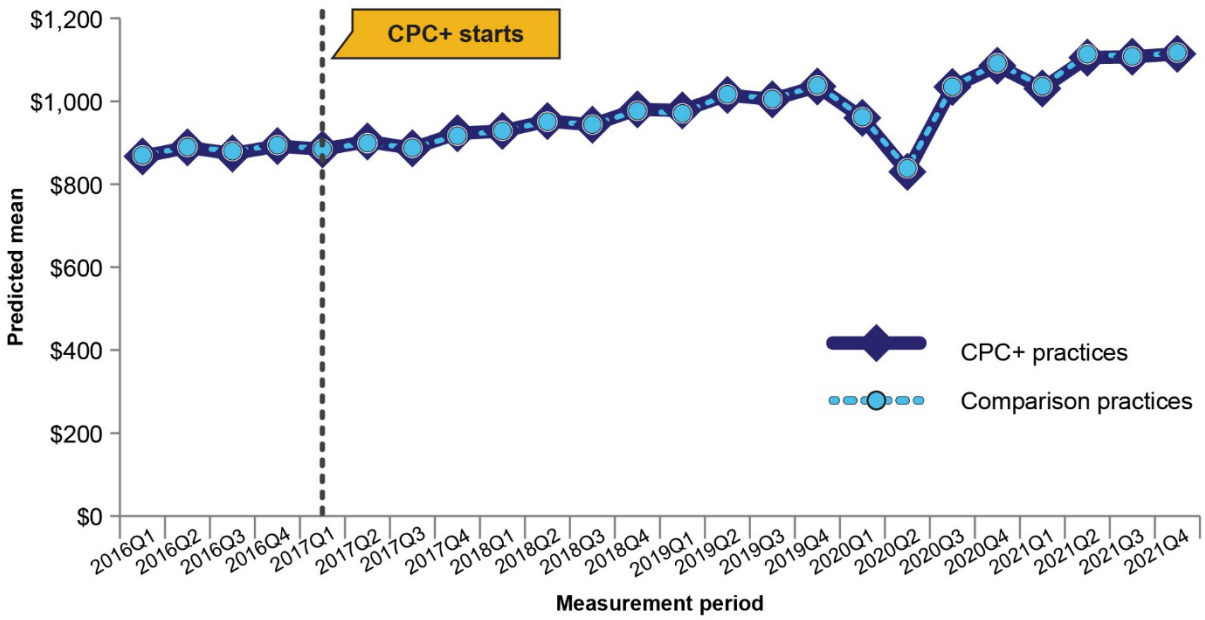
[m] After accounting for weights that adjust for matching and time observed in Medicare FFS, the effective sample sizes fall but are still substantial. For the comparison group, the effective sample size is 43 to 50 percent of the actual sample size. The effective sample size for the CPC+ group is 96 percent of the actual sample size because it is affected only by time observed (and not by the matching weights).

*/**/*** Significantly different from zero at the 0.10/0.05/0.01 level, two-tailed test.

NA = not applicable because the difference-in-differences impact estimate cannot be calculated at baseline.

ACO = Accountable Care Organization; APM = Alternative Payment Model; C = comparison; CAH = critical access hospital; ED = emergency department; FFS = fee-for-service; MIPS = Merit-based Incentive Payment System; NA = not applicable; PY = Program Year; QPP = Quality Payment Program; SE = standard error; SSP = Medicare Shared Savings Program.

Figure 5.A.2.1. Quarterly trends in average Medicare expenditures PBPM, excluding CMS's enhanced payments, Track 1

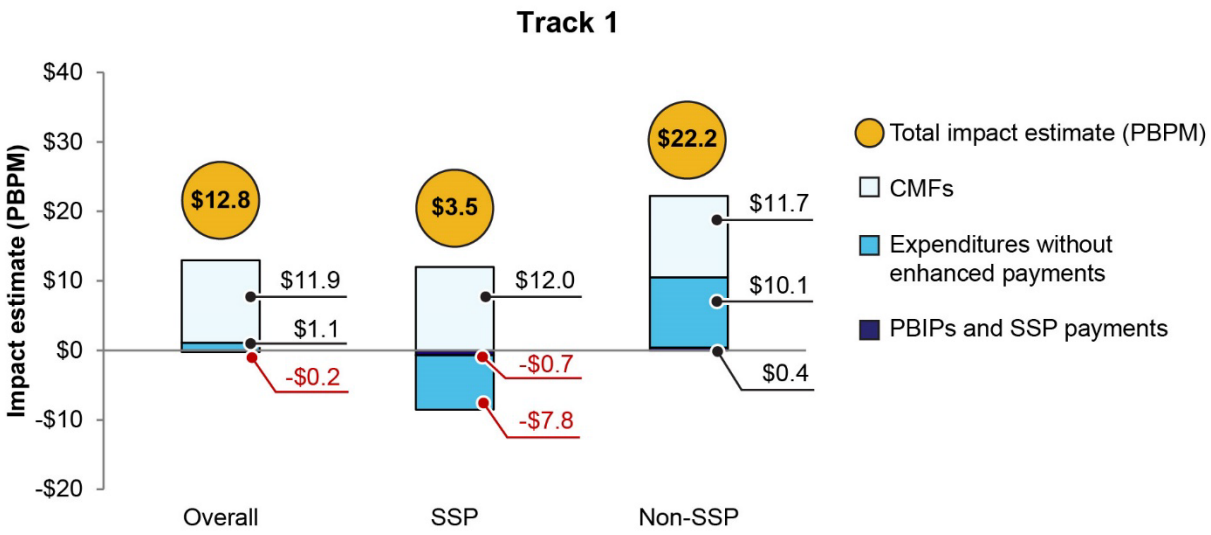


Source: Analyses of Medicare claims data from January 2013 through December 2021.

Notes: For beneficiaries attributed to CPC+ and comparison practices, the figure shows actual, unadjusted average expenditures in the baseline quarters (Q1 through Q4 of 2016), which are similar for the two groups due to matching. In the intervention quarters (starting in Q1 2017), the comparison group mean is regression-adjusted based on the quarterly difference-in-differences model, which controls for baseline characteristics and COVID-19 related controls. The sharp decline in expenditures during the first and second quarters of 2020 can be attributed to a decline in the overall utilization of health services during the initial months of the COVID-19 pandemic.

PBPM = per beneficiary per month.

Figure 5.A.2.2. Per beneficiary per month impact estimates for Medicare expenditures, with CMS's enhanced payments, by Track 1 and SSP status



Source: Mathematica's analysis of Medicare claims data from January 2013 through December 2021.

Notes: The impact estimates on expenditures without enhanced payments over the five years of CPC+ ($1.1 in Track 1 overall, -$7.8 in Track 1 SSP, $10.1 in Track 1 non-SSP) were not statistically significant overall but were statistically significant within SSP subgroups. The impact estimates on expenditures including enhanced payments that are attributable to PBIPs and SSP payments were smaller by $0.2 for Track 1 overall and by $0.7 for Track 1 SSP, compared to the respective impact estimates that do not include PBIPs and SSP payments. The estimates attributable to PBIPs and SSP payments are negative because, between the baseline and the intervention period, the change due to PBIPs were $13.7 and $4.4 higher for CPC+ practices than for comparison practices in Track 1 overall and Track 1 SSP, respectively (because only CPC+ practices receive PBIPs) and the change due to SSP payments were $13.9 and $5.1 lower for CPC+ practices than for comparison practices, for Track 1 overall and Track 1 SSP respectively. This resulted in the impact estimates decreasing by $0.2 and $0.7, for Track 1 overall and Track 1 SSP respectively, after including both PBIPs and SSP payments.

The figure includes: (1) 738 Track 1 SSP and 635 Track 1 non-SSP CPC+ practices that were participating in CPC+ as of April 1, 2017 (the end of the first program quarter), and (2) 2,979 Track 1 SSP and 2,264 Track 1 non-SSP comparison practices.

CMF = care management fee; CMS = Centers for Medicare and Medicaid Services; PBIP = Performance-based Incentive Payment; PBPM = per beneficiary per month; SSP = Medicare Shared Savings Program.

Table 5.A.2.2a. Regression-adjusted means and estimated impacts of CPC+ on selected Medicare expenditure outcomes for attributed Medicare FFS beneficiaries by program year and average across the five program years, Track 2

	Track 2—Overall					
	CPC+ mean[a]	C mean[a]	Impact estimate[b] (SE)	Percentage impact[c]	90 percent confidence interval	*p*-Value
Medicare expenditures (per beneficiary per month)						
Medicare Part A and B expenditures without enhanced payments for CPC+ and SSP[d]						
Baseline	$876	$877	NA	NA	NA	NA
PY 1	$897	$893	$5.1 ($3.5)	0.6%	(-$0.6, $10.8)	0.14
PY 2	$949	$945	$5.0 ($4.0)	0.5%	(-$1.6, $11.6)	0.22
PY 3	$989	$992	-$1.6 ($4.5)	-0.2%	(-$8.9, $5.8)	0.73
PY 4	$946	$949	-$2.3 ($5.1)	-0.2%	(-$10.8, $6.1)	0.65
PY 5	$1,034	$1,034	$0.7 ($5.6)	0.1%	(-$8.5, $9.9)	0.90
PY 1 through 5	$965	$965	$1.3 ($3.8)	0.1%	(-$5.0, $7.7)	0.73
Medicare Part A and B expenditures including care management fees						
Baseline	$876	$877	NA	NA	NA	NA
PY 1	$923	$893	$31.2*** ($3.5)	3.5%	($25.5, $37.0)	0.00
PY 2	$973	$945	$29.4*** ($4.0)	3.1%	($22.8, $36.0)	0.00
PY 3	$1,013	$992	$22.2*** ($4.5)	2.2%	($14.8, $29.5)	0.00
PY 4	$969	$950	$20.2*** ($5.2)	2.1%	($11.8, $28.7)	0.00
PY 5	$1,054	$1,034	$21.0*** ($5.6)	2.0%	($11.8, $30.2)	0.00
PY 1 through 5	$989	$965	$24.8*** ($3.9)	2.6%	($18.4, $31.1)	0.00
Medicare Part A and B expenditures including care management fees, Performance-based Incentive Payments, and shared savings payments to SSP ACOs[e]						
Baseline	$879	$880	NA	NA	NA	NA
PY 1	$925	$895	$31.2*** ($3.5)	3.5%	($25.5, $36.9)	0.00
PY 2	$976	$948	$29.8*** ($4.0)	3.1%	($23.2, $36.3)	0.00
PY 3	$1,017	$997	$21.5*** ($4.4)	2.2%	($14.2, $28.7)	0.00
PY 4	$976	$957	$19.5*** ($5.1)	2.0%	($11.1, $27.8)	0.00
PY 5	$1,060	$1,041	$20.0*** ($5.6)	1.9%	($10.9, $29.2)	0.00
PY 1 through 5	$993	$970	$24.4*** ($3.8)	2.5%	($18.1, $30.6)	0.00
Medicare expenditures by service category (per beneficiary per month)[f]						
Inpatient expenditures						
Baseline	$314	$317	NA	NA	NA	NA
PY 1	$321	$320	$3.5 ($2.3)	1.1%	(-$0.2, $7.3)	0.12
PY 2	$329	$329	$2.3 ($2.5)	0.7%	(-$1.8, $6.4)	0.36
PY 3	$336	$344	-$5.3** ($2.6)	-1.5%	(-$9.5, -$1.0)	0.04
PY 4	$320	$325	-$2.6 ($3.0)	-0.8%	(-$7.5, $2.3)	0.38

Table 5.A.2.2a. ***(continued)***

	Track 2—Overall					
	CPC+ mean[a]	C mean[a]	Impact estimate[b] (SE)	Percentage impact[c]	90 percent confidence interval	*p*-Value
PY 5	$336	$337	$1.3 ($3.1)	0.4%	(-$3.8, $6.4)	0.68
PY 1 through 5	$329	$331	-$0.2 ($2.2)	-0.1%	(-$3.9, $3.4)	0.92
Expenditures for acute inpatient care[g]						
Baseline	$278	$281	NA	NA	NA	NA
PY 1	$284	$285	$2.7 ($2.1)	1.0%	(-$0.6, $6.1)	0.18
PY 2	$292	$294	$0.7 ($2.2)	0.3%	(-$3.0, $4.4)	0.75
PY 3	$298	$308	-$6.8*** ($2.3)	-2.2%	(-$10.7, -$3.0)	0.00
PY 4	$284	$292	-$5.3** ($2.6)	-1.8%	(-$9.6, -$0.9)	0.04
PY 5	$298	$302	-$1.8 ($2.8)	-0.6%	(-$6.5, $2.8)	0.51
PY 1 through 5	$291	$297	-$2.2 ($2.0)	-0.7%	(-$5.5, $1.1)	0.28
Expenditures for acute surgical hospitalizations						
Baseline	$147	$148	NA	NA	NA	NA
PY 1	$149	$146	$4.1*** ($1.5)	2.8%	($1.6, $6.6)	0.01
PY 2	$151	$149	$3.4** ($1.6)	2.3%	($0.7, $6.0)	0.04
PY 3	$155	$157	-$0.5 ($1.5)	-0.3%	(-$3.1, $2.0)	0.73
PY 4	$140	$141	$0.8 ($1.7)	0.6%	(-$2.0, $3.6)	0.65
PY 5	$143	$144	$0.7 ($1.7)	0.5%	(-$2.2, $3.6)	0.69
PY 1 through 5	$147	$147	$1.6 ($1.3)	1.1%	(-$0.5, $3.8)	0.20
Expenditures for acute medical hospitalizations						
Baseline	$131	$133	NA	NA	NA	NA
PY 1	$135	$138	-$1.4 ($1.1)	-1.0%	(-$3.2, $0.5)	0.22
PY 2	$141	$145	-$2.6** ($1.2)	-1.8%	(-$4.7, -$0.6)	0.03
PY 3	$143	$151	-$6.3*** ($1.4)	-4.2%	(-$8.6, -$4.0)	0.00
PY 4	$144	$151	-$6.0*** ($1.5)	-4.0%	(-$8.6, -$3.5)	0.00
PY 5	$154	$159	-$2.5 ($1.7)	-1.6%	(-$5.4, $0.3)	0.14
PY 1 through 5	$144	$149	-$3.8*** ($1.2)	-2.6%	(-$5.8, -$1.9)	0.00
Inpatient rehabilitation facility expenditures						
Baseline	$20	$20	NA	NA	NA	NA
PY 1	$22	$21	$0.8* ($0.4)	4.0%	($0.1, $1.5)	0.05
PY 2	$22	$22	$1.3*** ($0.5)	6.3%	($0.5, $2.1)	0.01
PY 3	$23	$22	$1.3*** ($0.5)	6.1%	($0.5, $2.2)	0.01
PY 4	$23	$21	$2.3*** ($0.6)	10.9%	($1.3, $3.2)	0.00
PY 5	$25	$23	$2.4*** ($0.6)	10.3%	($1.4, $3.3)	0.00
PY 1 through 5	$23	$22	$1.6*** ($0.4)	7.5%	($1.0, $2.3)	0.00

Table 5.A.2.2a. ***(continued)***

	Track 2—Overall					
	CPC+ mean[a]	C mean[a]	Impact estimate[b] (SE)	Percentage impact[c]	90 percent confidence interval	*p*-Value
Post-acute care expenditures[h]						
Baseline	$109	$108	NA	NA	NA	NA
PY 1	$107	$106	$0.9 ($1.2)	0.8%	(-$1.1, $2.8)	0.46
PY 2	$108	$105	$1.5 ($1.2)	1.4%	(-$0.5, $3.4)	0.21
PY 3	$107	$105	$1.1 ($1.3)	1.0%	(-$1.0, $3.2)	0.39
PY 4	$96	$94	$1.6 ($1.4)	1.7%	(-$0.6, $3.9)	0.23
PY 5	$94	$90	$3.5** ($1.4)	3.8%	($1.2, $5.7)	0.01
PY 1 through 5	$102	$100	$1.7 ($1.0)	1.7%	($0.0, $3.4)	0.10
Acute inpatient and post-acute care expenditures combined[h]						
Baseline	$387	$389	NA	NA	NA	NA
PY 1	$392	$390	$3.6 ($2.8)	0.9%	(-$1.0, $8.2)	0.19
PY 2	$400	$399	$2.2 ($2.9)	0.6%	(-$2.5, $7.0)	0.44
PY 3	$405	$413	-$5.7* ($3.1)	-1.4%	(-$10.8, -$0.7)	0.06
PY 4	$380	$386	-$3.6 ($3.5)	-0.9%	(-$9.3, $2.1)	0.30
PY 5	$392	$392	$1.6 ($3.6)	0.4%	(-$4.4, $7.6)	0.66
PY 1 through 5	$394	$396	-$0.5 ($2.6)	-0.1%	(-$4.8, $3.8)	0.86
Outpatient expenditures						
Baseline	$166	$170	NA	NA	NA	NA
PY 1	$178	$181	$0.8 ($0.9)	0.5%	(-$0.6, $2.2)	0.35
PY 2	$199	$203	$0.1 ($1.2)	0.0%	(-$1.9, $2.1)	0.94
PY 3	$214	$219	-$1.4 ($1.7)	-0.7%	(-$4.1, $1.3)	0.40
PY 4	$204	$212	-$4.7** ($2.0)	-2.3%	(-$8.0, -$1.5)	0.02
PY 5	$230	$242	-$7.5*** ($2.1)	-3.2%	(-$11.0, -$4.0)	0.00
PY 1 through 5	$206	$213	-$2.5* ($1.3)	-1.2%	(-$4.7, -$0.3)	0.06
Expenditures for outpatient ED visits, including observation stays[i]						
Baseline	$25	$26	NA	NA	NA	NA
PY 1	$27	$27	-$0.1 ($0.2)	-0.2%	(-$0.4, $0.2)	0.75
PY 2	$28	$29	-$0.2 ($0.2)	-0.9%	(-$0.6, $0.1)	0.29
PY 3	$29	$30	-$0.3 ($0.3)	-1.0%	(-$0.7, $0.1)	0.24
PY 4	$24	$26	-$0.6** ($0.3)	-2.3%	(-$1.0, -$0.1)	0.05
PY 5	$28	$30	-$1.0*** ($0.4)	-3.5%	(-$1.6, -$0.4)	0.01
PY 1 through 5	$27	$28	-$0.4** ($0.2)	-1.6%	(-$0.8, -$0.1)	0.05

Table 5.A.2.2a. ***(continued)***

	Track 2—Overall					
	CPC+ mean[a]	C mean[a]	Impact estimate[b] (SE)	Percentage impact[c]	90 percent confidence interval	*p*-Value
Expenditures for physician and nonphysician Part B noninstitutional services in any setting[j]						
Baseline	$245	$239	NA	NA	NA	NA
PY 1	$251	$244	-$0.1 ($0.8)	0.0%	(-$1.3, $1.2)	0.91
PY 2	$265	$259	$0.1 ($1.1)	0.0%	(-$1.7, $1.9)	0.95
PY 3	$278	$271	$0.8 ($1.4)	0.3%	(-$1.5, $3.1)	0.59
PY 4	$262	$252	$3.0* ($1.6)	1.1%	($0.3, $5.6)	0.06
PY 5	$302	$293	$2.4 ($1.8)	0.8%	(-$0.6, $5.5)	0.19
PY 1 through 5	$272	$265	$1.2 ($1.2)	0.4%	(-$0.7, $3.1)	0.31
Expenditures for ambulatory visits with primary care practitioners						
Baseline	$24	$24	NA	NA	NA	NA
PY 1	$25	$25	$0.1 ($0.1)	0.4%	(-$0.1, $0.3)	0.32
PY 2	$27	$26	$1.0*** ($0.1)	3.9%	($0.8, $1.2)	0.00
PY 3	$28	$27	$1.2*** ($0.2)	4.5%	($0.9, $1.5)	0.00
PY 4	$26	$24	$2.2*** ($0.2)	9.2%	($1.9, $2.5)	0.00
PY 5	$33	$31	$1.6*** ($0.2)	5.0%	($1.2, $2.0)	0.00
PY 1 through 5	$28	$27	$1.2*** ($0.1)	4.6%	($1.0, $1.5)	0.00
Expenditures for ambulatory visits with primary care practitioners at assigned practice[k]						
Baseline	$17	$17	NA	NA	NA	NA
PY 1	$17	$17	$0.3*** ($0.1)	1.8%	($0.1, $0.5)	0.00
PY 2	$17	$15	$1.6*** ($0.1)	10.2%	($1.3, $1.8)	0.00
PY 3	$17	$15	$2.0*** ($0.2)	12.9%	($1.7, $2.3)	0.00
PY 4	$15	$12	$2.8*** ($0.2)	23.2%	($2.5, $3.2)	0.00
PY 5	$17	$14	$2.7*** ($0.3)	18.6%	($2.3, $3.2)	0.00
PY 1 through 5	$17	$14	$1.9*** ($0.2)	12.7%	($1.6, $2.1)	0.00
Expenditures for ambulatory visits with primary care practitioners at non-assigned practice[k]						
Baseline	$7	$8	NA	NA	NA	NA
PY 1	$8	$9	-$0.2*** ($0.1)	-2.7%	(-$0.3, -$0.1)	0.00
PY 2	$10	$11	-$0.6*** ($0.1)	-5.5%	(-$0.7, -$0.4)	0.00
PY 3	$11	$12	-$0.8*** ($0.1)	-6.5%	(-$1.0, -$0.6)	0.00
PY 4	$11	$13	-$0.6*** ($0.2)	-5.1%	(-$0.9, -$0.3)	0.00
PY 5	$15	$17	-$1.1*** ($0.2)	-7.0%	(-$1.5, -$0.8)	0.00
PY 1 through 5	$11	$13	-$0.7*** ($0.1)	-5.6%	(-$0.8, -$0.5)	0.00
Expenditures for ambulatory visits with specialists						
Baseline	$24	$24	NA	NA	NA	NA
PY 1	$24	$23	$0.0 ($0.1)	-0.2%	(-$0.1, $0.1)	0.50

Table 5.A.2.2a. ***(continued)***

	Track 2—Overall					
	CPC+ mean[a]	C mean[a]	Impact estimate[b] (SE)	Percentage impact[c]	90 percent confidence interval	*p*-Value
PY 2	$24	$24	-$0.1 ($0.1)	-0.5%	(-$0.3, $0.0)	0.13
PY 3	$24	$24	-$0.2** ($0.1)	-0.9%	(-$0.4, -$0.1)	0.03
PY 4	$21	$21	-$0.2** ($0.1)	-1.2%	(-$0.4, -$0.1)	0.03
PY 5	$28	$28	-$0.3** ($0.2)	-1.2%	(-$0.6, -$0.1)	0.03
PY 1 through 5	$24	$24	-$0.2** ($0.1)	-0.8%	(-$0.3, $0.0)	0.03
Expenditures on laboratory services[l]						
Baseline	$23	$23	NA	NA	NA	NA
PY 1	$24	$24	-$0.3** ($0.1)	-1.1%	(-$0.5, -$0.1)	0.01
PY 2	$26	$26	-$0.3* ($0.2)	-1.3%	(-$0.6, $0.0)	0.06
PY 3	$26	$26	$0.0 ($0.2)	-0.2%	(-$0.3, $0.2)	0.78
PY 4	$28	$28	-$0.3 ($0.2)	-1.0%	(-$0.6, $0.0)	0.16
PY 5	$32	$33	-$1.2*** ($0.2)	-3.6%	(-$1.6, -$0.8)	0.00
PY 1 through 5	$27	$28	-$0.4*** ($0.1)	-1.5%	(-$0.7, -$0.2)	0.00
Expenditures on imaging services[l]						
Baseline	$45	$45	NA	NA	NA	NA
PY 1	$46	$46	-$0.1 ($0.2)	-0.3%	(-$0.5, $0.3)	0.63
PY 2	$49	$49	-$0.2 ($0.3)	-0.4%	(-$0.6, $0.2)	0.46
PY 3	$52	$52	-$0.3 ($0.3)	-0.5%	(-$0.7, $0.2)	0.38
PY 4	$46	$47	-$0.6* ($0.3)	-1.2%	(-$1.1, $0.0)	0.09
PY 5	$52	$53	-$0.9** ($0.4)	-1.6%	(-$1.5, -$0.3)	0.02
PY 1 through 5	$49	$50	-$0.4* ($0.2)	-0.8%	(-$0.8, $0.0)	0.10
Skilled nursing facility expenditures						
Baseline	$65	$64	NA	NA	NA	NA
PY 1	$63	$62	$0.1 ($0.7)	0.2%	(-$1.1, $1.3)	0.84
PY 2	$64	$63	$0.6 ($0.8)	0.9%	(-$0.7, $1.8)	0.45
PY 3	$63	$62	$0.5 ($0.9)	0.7%	(-$0.9, $1.9)	0.59
PY 4	$63	$62	$0.2 ($1.0)	0.3%	(-$1.4, $1.8)	0.83
PY 5	$65	$62	$2.4** ($1.1)	3.8%	($0.6, $4.1)	0.03
PY 1 through 5	$64	$62	$0.8 ($0.7)	1.2%	(-$0.4, $1.9)	0.30
Home health expenditures						
Baseline	$41	$41	NA	NA	NA	NA
PY 1	$40	$41	-$0.3 ($0.3)	-0.8%	(-$0.8, $0.2)	0.31
PY 2	$41	$42	-$0.7** ($0.4)	-1.8%	(-$1.3, -$0.2)	0.04
PY 3	$41	$42	-$0.9** ($0.4)	-2.2%	(-$1.6, -$0.3)	0.02

Table 5.A.2.2a. ***(continued)***

	Track 2—Overall					
	CPC+ mean[a]	C mean[a]	Impact estimate[b] (SE)	Percentage impact[c]	90 percent confidence interval	*p*-Value
PY 4	$37	$39	-$1.6*** ($0.4)	-4.2%	(-$2.4, -$0.9)	0.00
PY 5	$39	$41	-$2.0*** ($0.5)	-4.9%	(-$2.8, -$1.2)	0.00
PY 1 through 5	$40	$41	-$1.1*** ($0.3)	-2.8%	(-$1.7, -$0.6)	0.00
Hospice expenditures						
Baseline	$24	$25	NA	NA	NA	NA
PY 1	$24	$25	$0.7 ($0.4)	2.8%	($0.0, $1.4)	0.11
PY 2	$28	$27	$2.4*** ($0.5)	9.4%	($1.5, $3.3)	0.00
PY 3	$31	$30	$3.2*** ($0.6)	11.3%	($2.2, $4.2)	0.00
PY 4	$33	$32	$2.1*** ($0.6)	6.8%	($1.1, $3.1)	0.00
PY 5	$34	$33	$2.9*** ($0.7)	9.3%	($1.8, $4.1)	0.00
PY 1 through 5	$31	$30	$2.3*** ($0.5)	8.2%	($1.5, $3.1)	0.00
Durable medical equipment expenditures						
Baseline	$21	$21	NA	NA	NA	NA
PY 1	$20	$20	$0.3 ($0.2)	1.5%	(-$0.1, $0.7)	0.23
PY 2	$23	$22	$0.3 ($0.3)	1.2%	(-$0.2, $0.7)	0.35
PY 3	$24	$24	$0.5 ($0.3)	2.2%	($0.0, $1.0)	0.11
PY 4	$26	$25	$0.6* ($0.3)	2.4%	($0.0, $1.2)	0.08
PY 5	$26	$26	$0.3 ($0.3)	1.0%	(-$0.3, $0.8)	0.44
PY 1 through 5	$24	$23	$0.4 ($0.3)	1.7%	($0.0, $0.8)	0.14
Unweighted sample sizes[m]						
Number of practices	1,515	3,783				
Number of beneficiaries	1,896,880	4,507,499				
Number of beneficiary-years	7,225,289	17,054,519				

Source: Mathematica's analysis of Medicare claims data from January 2013 through December 2021.

Notes: Although this table indicates which estimates are statistically significant, when we interpret evidence, we combine evidence from the magnitude of the effect, the *p*-values, findings on related outcomes, subgroups, sensitivity tests, and other data sources about model implementation.

[a] We report the actual, unadjusted averages in the baseline period which are similar for the CPC+ and comparison groups due to matching. In the intervention periods, the comparison group mean is computed by subtracting the regression adjusted difference between the CPC+ and comparison means in each time period from the CPC+ mean in that same time period.

[b] Each impact estimate is regression-adjusted using a difference-in-differences analysis that reflects the difference of the average outcome for Medicare FFS beneficiaries attributed to CPC+ practices in the first five years of CPC+ to the average outcome in the baseline year, relative to the same difference over time for Medicare FFS beneficiaries attributed to comparison practices, while controlling for beneficiary characteristics, practice fixed effects and COVID-19 related controls.

[c] We calculated percentage impacts relative to what the CPC+ mean would have been in Program Years 1 through 5 (separately and combined) in the absence of the intervention—that is, the unadjusted CPC+ mean minus the impact estimate.

[d] Expenditures for Part A and Part B services in PY 3, PY 4, and PY 5 include QPP payment adjustments, based on practitioner performance two years before. They are applicable for both CPC+ and comparison practices. The adjustments are composed of (1) MIPS adjustments, which are applied directly to physician and outpatient claims (as a percentage of the charges on the claims); and (2) lump sum incentive payments to eligible practitioners who participated in Advanced APMs in 2017, 2018, and 2019 (calculated based on 2018, 2019, and 2020 claims for these practitioners, respectively). The first QPP adjustments were paid in PY 3 (two years after the start of QPP), so there are no QPP payments in PYs 1 and 2.

Table 5.A.2.2a. ***(continued)***

[e] We determine SSP ACO participation status based on participation at the beginning of PY 1 (January 1, 2017). However, over time, CPC+ practices may join or leave SSP, resulting in a small subset of SSP practices receiving the Performance-based Incentive Payments and a small subset of non-SSP practices receiving the shared savings payments. This is reflected in the impact estimates.

[f] The sum of expenditures by claim type does not equal the total expenditures for Part A and B services without enhanced payments in PY 3, PY 4, and PY 5 because the total expenditures include lum*p*-sum incentive payments that are not applied at the claim level and are instead paid out directly to eligible practitioners who participated in Advanced APMs in 2017, 2018 and 2019.

[g] Acute inpatient care includes short-stay acute hospital admissions and admissions to CAHs. Expenditures for non-acute hospital admissions other than those for inpatient rehabilitation, such as psychiatric hospital admissions, are included in inpatient expenditures but not shown separately.

[h] Post-acute care expenditures include expenditures on home health, long-term care, skilled nursing facility, and inpatient rehabilitation. These are not a sub-category of inpatient expenditures.

[i] Expenditures, with QPP payment adjustments, on outpatient ED visits include professional (which is part of expenditures for physician and nonphysician Part B noninstitutional services) and facility fees, as well as payments for observation stays.

[j] Expenditures, with QPP payment adjustments, on Part B noninstitutional services include expenditures for (1) ambulatory primary care visits, (2) ambulatory specialist visits, and (3) non-ambulatory physician visits as well as services provided by other noninstitutional providers. (We only show the first two categories separately in the table).

[k] We define the assigned practice for the baseline period as the first practice to which a beneficiary was attributed during the baseline period, and the assigned practice for the intervention period as the first practice that the beneficiary was attributed to during the intervention period.

[l] Laboratory and imaging services were identified in both the carrier and outpatient claim files.

[m] After accounting for weights that adjust for matching and time observed in Medicare FFS, the effective sample sizes fall but are still substantial. For the comparison group, the effective sample size is 40 percent of the actual sample size. The effective sample size for the CPC+ group is 96 percent of the actual sample size because it is affected only by time observed (and not by the matching weights).

*/**/*** Significantly different from zero at the 0.10/0.05/0.01 level, two-tailed test.

NA = not applicable because the difference-in-differences impact estimate cannot be calculated at baseline.

ACO = Accountable Care Organization; APM = Alternative Payment Model; C = comparison; CAH = critical access hospital; ED = emergency department; FFS = fee-for-service; MIPS = Merit-based Incentive Payment System; NA = not applicable; PY = Program Year; QPP = Quality Payment Program; SE = standard error.

Table 5.A.2.2b. Regression-adjusted means and estimated impacts of CPC+ on selected Medicare expenditure outcomes for attributed Medicare FFS beneficiaries, by program year, and average across the five program years, Track 2 by SSP status

	Track 2—SSP						Track 2—Non-SSP						
	CPC+ mean[a]	C mean[a]	Impact estimate[b] (SE)	Percentage impact[c]	90 percent confidence interval	*p*-Value	CPC+ mean[a]	C mean[a]	Impact estimate[b] (SE)	Percentage impact[c]	90 percent confidence interval	*p*-Value	*p*-Value for SSP vs. non-SSP difference
Medicare expenditures (per beneficiary per month)													
Medicare Part A and B expenditures without enhanced payments for CPC+ and SSP[d]													
Baseline	$896	$893	NA	NA	NA	NA	$861	$865	NA	NA	NA	NA	NA
PY 1	$917	$913	$1.3 ($5.1)	0.1%	(-$7.1, $9.8)	0.80	$881	$877	$7.9* ($4.7)	0.9%	($0.2, $15.6)	0.09	0.34
PY 2	$966	$966	-$2.5 ($6.2)	-0.3%	(-$12.7, $7.7)	0.69	$935	$928	$10.8** ($5.2)	1.2%	($2.2, $19.4)	0.04	0.10
PY 3	$1,009	$1,014	-$8.0 ($7.1)	-0.8%	(-$19.7, $3.7)	0.26	$974	$975	$3.3 ($5.7)	0.3%	(-$6.0, $12.6)	0.55	0.21
PY 4	$956	$968	-$14.3* ($8.2)	-1.5%	(-$27.8, -$0.9)	0.08	$938	$933	$9.0 ($6.0)	1.0%	(-$0.9, $18.8)	0.14	0.02
PY 5	$1,048	$1,063	-$17.4* ($9.1)	-1.6%	(-$32.4, -$2.5)	0.06	$1,022	$1,012	$14.1** ($6.7)	1.4%	($3.2, $25.1)	0.03	0.01
PY 1 through 5	$982	$987	-$8.1 ($6.1)	-0.8%	(-$18.1, $1.9)	0.18	$952	$948	$8.9* ($4.7)	0.9%	($1.2, $16.6)	0.06	0.03
Medicare Part A and B expenditures including care management fees													
Baseline	$896	$893	NA	NA	NA	NA	$861	$865	NA	NA	NA	NA	NA
PY 1	$943	$913	$27.4*** ($5.2)	3.0%	($19.0, $35.9)	0.00	$907	$877	$34.2*** ($4.7)	3.9%	($26.5, $41.9)	0.00	0.34
PY 2	$990	$966	$21.4*** ($6.2)	2.2%	($11.2, $31.6)	0.00	$959	$928	$35.6*** ($5.2)	3.9%	($27.0, $44.3)	0.00	0.08
PY 3	$1,032	$1,014	$15.4** ($7.1)	1.5%	($3.6, $27.1)	0.03	$998	$975	$27.4*** ($5.7)	2.8%	($18.1, $36.7)	0.00	0.19
PY 4	$979	$968	$8.0 ($8.2)	0.8%	(-$5.5, $21.5)	0.33	$961	$933	$31.7*** ($6.0)	3.4%	($21.8, $41.6)	0.00	0.02
PY 5	$1,069	$1,063	$2.8 ($9.1)	0.3%	(-$12.1, $17.8)	0.76	$1,043	$1,012	$34.5*** ($6.7)	3.4%	($23.5, $45.5)	0.00	0.00
PY 1 through 5	$1,005	$987	$15.1** ($6.1)	1.5%	($5.1, $25.2)	0.01	$976	$947	$32.6*** ($4.7)	3.5%	($24.8, $40.3)	0.00	0.02
Medicare Part A and B expenditures including care management fees, Performance-based Incentive Payments, and shared savings payments to SSP ACOs[e]													
Baseline	$901	$899	NA	NA	NA	NA	$861	$865	NA	NA	NA	NA	NA
PY 1	$946	$919	$24.7*** ($5.1)	2.7%	($16.3, $33.1)	0.00	$909	$877	$36.2*** ($4.7)	4.1%	($28.5, $43.9)	0.00	0.10
PY 2	$994	$972	$20.1*** ($6.1)	2.1%	($10.2, $30.1)	0.00	$962	$928	$37.3*** ($5.2)	4.0%	($28.7, $45.9)	0.00	0.03
PY 3	$1,037	$1,022	$13.1* ($7.0)	1.3%	($1.6, $24.5)	0.06	$1,001	$977	$28.0*** ($5.6)	2.9%	($18.7, $37.2)	0.00	0.10
PY 4	$989	$980	$6.2 ($8.0)	0.6%	(-$7.0, $19.5)	0.44	$965	$937	$31.8*** ($6.0)	3.4%	($22.0, $41.7)	0.00	0.01

Table 5.A.2.2b. ***(continued)***

	Track 2—SSP						Track 2—Non-SSP						
	CPC+ mean[a]	C mean[a]	Impact estimate[b] (SE)	Percentage impact[c]	90 percent confidence interval	*p*-Value	CPC+ mean[a]	C mean[a]	Impact estimate[b] (SE)	Percentage impact[c]	90 percent confidence interval	*p*-Value	*p*-Value for SSP vs. non-SSP difference
PY 5	$1,076	$1,073	$0.7 ($9.0)	0.1%	(-$14.2, $15.6)	0.94	$1,047	$1,016	$34.4*** ($6.7)	3.4%	($23.4, $45.3)	0.00	0.00
PY 1 through 5	$1,011	$995	$13.1** ($6.0)	1.3%	($3.3, $22.9)	0.03	$980	$950	$33.5*** ($4.7)	3.5%	($25.7, $41.2)	0.00	0.01
Medicare expenditures by service category (per beneficiary per month)[f]													
Inpatient expenditures													
Baseline	$322	$322	NA	NA	NA	NA	$308	$312	NA	NA	NA	NA	NA
PY 1	$330	$329	$0.6 ($3.4)	0.2%	(-$5.0, $6.2)	0.86	$314	$312	$5.9* ($3.1)	1.9%	($0.8, $11.0)	0.06	0.25
PY 2	$335	$337	-$2.0 ($3.8)	-0.6%	(-$8.3, $4.3)	0.60	$324	$322	$5.7* ($3.3)	1.8%	($0.3, $11.1)	0.08	0.13
PY 3	$344	$351	-$7.2* ($4.0)	-2.1%	(-$13.8, -$0.7)	0.07	$330	$338	-$3.8 ($3.4)	-1.1%	(-$9.4, $1.9)	0.27	0.51
PY 4	$327	$333	-$6.9 ($4.6)	-2.1%	(-$14.6, $0.7)	0.14	$315	$318	$2.3 ($3.5)	0.8%	(-$3.5, $8.2)	0.51	0.11
PY 5	$343	$349	-$5.6 ($5.0)	-1.6%	(-$13.9, $2.6)	0.26	$330	$328	$6.2* ($3.8)	1.9%	($0.0, $12.4)	0.10	0.06
PY 1 through 5	$336	$340	-$4.3 ($3.4)	-1.3%	(-$9.9, $1.4)	0.21	$323	$324	$3.1 ($2.8)	1.0%	(-$1.5, $7.7)	0.26	0.09
Expenditures for acute inpatient care[g]													
Baseline	$286	$285	NA	NA	NA	NA	$271	$278	NA	NA	NA	NA	NA
PY 1	$293	$292	-$0.6 ($3.1)	-0.2%	(-$5.7, $4.5)	0.85	$278	$279	$5.4** ($2.7)	2.0%	($0.9, $9.9)	0.05	0.15
PY 2	$298	$300	-$3.5 ($3.5)	-1.2%	(-$9.3, $2.3)	0.32	$287	$289	$4.1 ($2.9)	1.4%	(-$0.7, $8.8)	0.16	0.09
PY 3	$306	$314	-$9.0** ($3.6)	-2.9%	(-$15.0, -$3.1)	0.01	$292	$303	-$5.2* ($3.1)	-1.7%	(-$10.2, -$0.2)	0.09	0.42
PY 4	$291	$298	-$8.8** ($4.1)	-2.9%	(-$15.5, -$2.1)	0.03	$279	$286	-$0.9 ($3.2)	-0.3%	(-$6.1, $4.3)	0.77	0.13
PY 5	$305	$312	-$8.0* ($4.6)	-2.6%	(-$15.5, -$0.5)	0.08	$291	$295	$3.0 ($3.4)	1.0%	(-$2.5, $8.6)	0.37	0.05
PY 1 through 5	$299	$304	-$6.0* ($3.2)	-2.0%	(-$11.2, -$0.8)	0.06	$285	$291	$1.2 ($2.5)	0.4%	(-$2.9, $5.3)	0.64	0.07
Expenditures for acute surgical hospitalizations													
Baseline	$151	$150	NA	NA	NA	NA	$143	$146	NA	NA	NA	NA	NA
PY 1	$152	$151	$0.5 ($2.3)	0.3%	(-$3.3, $4.3)	0.82	$147	$143	$7.0*** ($2.0)	5.0%	($3.7, $10.2)	0.00	0.04
PY 2	$152	$154	-$1.9 ($2.5)	-1.2%	(-$6.1, $2.3)	0.45	$150	$145	$7.5*** ($2.0)	5.3%	($4.2, $10.9)	0.00	0.00
PY 3	$158	$161	-$3.4 ($2.3)	-2.1%	(-$7.1, $0.3)	0.14	$152	$154	$1.7 ($2.1)	1.1%	(-$1.8, $5.1)	0.43	0.10

Table 5.A.2.2b. ***(continued)***

	Track 2—SSP						Track 2—Non-SSP						
	CPC+ mean[a]	C mean[a]	Impact estimate[b] (SE)	Percentage impact[c]	90 percent confidence interval	*p*-Value	CPC+ mean[a]	C mean[a]	Impact estimate[b] (SE)	Percentage impact[c]	90 percent confidence interval	*p*-Value	*p*-Value for SSP vs. non-SSP difference
PY 4	$142	$145	-$3.1 ($2.6)	-2.1%	(-$7.4, $1.1)	0.23	$138	$136	$4.9** ($2.2)	3.7%	($1.3, $8.5)	0.02	0.02
PY 5	$145	$147	-$3.4 ($2.7)	-2.3%	(-$7.8, $1.1)	0.21	$142	$140	$4.7** ($2.3)	3.4%	($0.9, $8.5)	0.04	0.02
PY 1 through 5	$150	$151	-$2.3 ($2.0)	-1.5%	(-$5.5, $1.0)	0.25	$146	$144	$5.1*** ($1.7)	3.6%	($2.3, $7.9)	0.00	0.00
Expenditures for acute medical hospitalizations													
Baseline	$135	$134	NA	NA	NA	NA	$128	$131	NA	NA	NA	NA	NA
PY 1	$141	$141	-$1.1 ($1.7)	-0.8%	(-$4.0, $1.7)	0.51	$131	$136	-$1.5 ($1.5)	-1.2%	(-$4.0, $0.9)	0.30	0.86
PY 2	$146	$147	-$1.6 ($2.0)	-1.1%	(-$4.8, $1.6)	0.41	$137	$143	-$3.4** ($1.6)	-2.5%	(-$6.1, -$0.8)	0.03	0.47
PY 3	$148	$153	-$5.6** ($2.2)	-3.7%	(-$9.3, -$1.9)	0.01	$140	$150	-$6.8*** ($1.7)	-4.7%	(-$9.7, -$4.0)	0.00	0.66
PY 4	$148	$153	-$5.7** ($2.5)	-3.7%	(-$9.8, -$1.6)	0.02	$140	$149	-$5.8*** ($1.8)	-4.0%	(-$8.8, -$2.8)	0.00	0.95
PY 5	$161	$165	-$4.6 ($2.9)	-2.8%	(-$9.4, $0.1)	0.11	$150	$154	-$1.7 ($1.9)	-1.1%	(-$4.8, $1.5)	0.39	0.39
PY 1 through 5	$149	$152	-$3.7* ($1.9)	-2.5%	(-$6.9, -$0.6)	0.05	$140	$147	-$3.9*** ($1.4)	-2.7%	(-$6.2, -$1.6)	0.00	0.93
Inpatient rehabilitation facility expenditures													
Baseline	$20	$22	NA	NA	NA	NA	$20	$20	NA	NA	NA	NA	NA
PY 1	$22	$22	$0.9 ($0.6)	4.2%	(-$0.1, $1.9)	0.16	$21	$20	$0.8 ($0.6)	3.8%	(-$0.2, $1.7)	0.17	0.88
PY 2	$23	$23	$0.8 ($0.8)	3.5%	(-$0.5, $2.0)	0.31	$22	$21	$1.8*** ($0.6)	8.6%	($0.7, $2.8)	0.01	0.31
PY 3	$23	$23	$0.8 ($0.8)	3.6%	(-$0.5, $2.1)	0.31	$23	$21	$1.7** ($0.7)	8.1%	($0.6, $2.9)	0.01	0.36
PY 4	$23	$22	$1.5 ($0.9)	6.8%	($0.0, $2.9)	0.10	$23	$20	$2.9*** ($0.7)	14.0%	($1.7, $4.1)	0.00	0.22
PY 5	$25	$24	$2.1** ($0.9)	8.9%	($0.6, $3.5)	0.02	$25	$23	$2.1*** ($0.7)	9.3%	($0.9, $3.3)	0.00	0.91
PY 1 through 5	$23	$23	$1.2* ($0.6)	5.3%	($0.1, $2.2)	0.06	$23	$21	$1.8*** ($0.5)	8.7%	($1.0, $2.7)	0.00	0.41
Post-acute care expenditures[h]													
Baseline	$113	$115	NA	NA	NA	NA	$106	$102	NA	NA	NA	NA	NA
PY 1	$112	$112	$2.1 ($1.8)	1.9%	(-$0.8, $5.0)	0.23	$104	$101	-$0.2 ($1.6)	-0.2%	(-$2.8, $2.4)	0.91	0.33
PY 2	$112	$111	$2.6 ($1.8)	2.4%	(-$0.3, $5.6)	0.14	$104	$100	$0.5 ($1.6)	0.5%	(-$2.1, $3.1)	0.74	0.38
PY 3	$111	$111	$1.9 ($1.9)	1.8%	(-$1.2, $5.1)	0.31	$104	$100	$0.4 ($1.8)	0.4%	(-$2.5, $3.3)	0.83	0.55

Table 5.A.2.2b. ***(continued)***

	Track 2—SSP						Track 2—Non-SSP						
	CPC+ mean[a]	C mean[a]	Impact estimate[b] (SE)	Percentage impact[c]	90 percent confidence interval	*p*-Value	CPC+ mean[a]	C mean[a]	Impact estimate[b] (SE)	Percentage impact[c]	90 percent confidence interval	*p*-Value	*p*-Value for SSP vs. non-SSP difference
PY 4	$99	$99	$1.9 ($2.1)	1.9%	(-$1.7, $5.4)	0.38	$94	$89	$1.8 ($1.7)	2.0%	(-$1.0, $4.6)	0.29	0.99
PY 5	$97	$96	$3.2 ($2.1)	3.4%	(-$0.2, $6.6)	0.12	$92	$86	$2.4 ($1.7)	2.7%	(-$0.3, $5.2)	0.15	0.77
PY 1 through 5	$106	$105	$2.3 ($1.6)	2.3%	(-$0.3, $5.0)	0.14	$99	$95	$1.0 ($1.3)	1.0%	(-$1.3, $3.2)	0.48	0.51
Acute inpatient and post-acute care expenditures combined[h]													
Baseline	$399	$400	NA	NA	NA	NA	$377	$380	NA	NA	NA	NA	NA
PY 1	$405	$404	$1.5 ($4.1)	0.4%	(-$5.3, $8.4)	0.71	$382	$379	$5.2 ($3.7)	1.4%	(-$0.9, $11.3)	0.16	0.51
PY 2	$410	$412	-$0.9 ($4.4)	-0.2%	(-$8.1, $6.3)	0.84	$391	$389	$4.6 ($3.8)	1.2%	(-$1.7, $10.9)	0.23	0.35
PY 3	$417	$425	-$7.1 ($4.6)	-1.7%	(-$14.7, $0.5)	0.13	$396	$403	-$4.8 ($4.1)	-1.2%	(-$11.5, $1.9)	0.24	0.71
PY 4	$389	$397	-$7.0 ($5.5)	-1.8%	(-$15.9, $2.0)	0.20	$373	$375	$0.9 ($4.1)	0.2%	(-$5.9, $7.7)	0.83	0.25
PY 5	$402	$408	-$4.8 ($5.9)	-1.2%	(-$14.5, $4.9)	0.42	$384	$381	$5.5 ($4.3)	1.4%	(-$1.6, $12.5)	0.20	0.16
PY 1 through 5	$405	$409	-$3.7 ($4.1)	-0.9%	(-$10.4, $3.0)	0.36	$385	$386	$2.1 ($3.3)	0.6%	(-$3.3, $7.5)	0.52	0.27
Outpatient expenditures													
Baseline	$175	$166	NA	NA	NA	NA	$160	$173	NA	NA	NA	NA	NA
PY 1	$187	$177	$1.5 ($1.3)	0.8%	(-$0.7, $3.7)	0.27	$171	$184	$0.3 ($1.1)	0.2%	(-$1.6, $2.1)	0.80	0.49
PY 2	$209	$200	$0.5 ($2.0)	0.2%	(-$2.8, $3.8)	0.81	$192	$205	-$0.2 ($1.5)	-0.1%	(-$2.7, $2.2)	0.88	0.78
PY 3	$225	$218	-$1.8 ($2.9)	-0.8%	(-$6.6, $3.0)	0.54	$205	$220	-$1.1 ($1.9)	-0.5%	(-$4.1, $2.0)	0.56	0.84
PY 4	$211	$209	-$6.5** ($3.3)	-3.0%	(-$11.9, -$1.1)	0.05	$198	$214	-$3.1 ($2.1)	-1.5%	(-$6.5, $0.4)	0.15	0.38
PY 5	$239	$241	-$9.7*** ($3.5)	-3.9%	(-$15.5, -$3.9)	0.01	$223	$243	-$5.7** ($2.4)	-2.5%	(-$9.6, -$1.7)	0.02	0.34
PY 1 through 5	$215	$210	-$3.1 ($2.3)	-1.4%	(-$6.9, $0.6)	0.17	$199	$214	-$1.9 ($1.5)	-0.9%	(-$4.3, $0.5)	0.20	0.65
Expenditures for outpatient ED visits, including observation stays[i]													
Baseline	$25	$27	NA	NA	NA	NA	$26	$26	NA	NA	NA	NA	NA
PY 1	$26	$28	-$0.3 ($0.3)	-1.2%	(-$0.8, $0.1)	0.24	$27	$27	$0.2 ($0.2)	0.6%	(-$0.2, $0.6)	0.53	0.20
PY 2	$28	$30	-$0.3 ($0.4)	-1.0%	(-$0.9, $0.3)	0.44	$29	$29	-$0.2 ($0.3)	-0.8%	(-$0.7, $0.3)	0.47	0.91
PY 3	$28	$31	-$0.8* ($0.4)	-2.8%	(-$1.6, -$0.1)	0.06	$30	$30	$0.1 ($0.3)	0.4%	(-$0.4, $0.6)	0.72	0.08

Table 5.A.2.2b. ***(continued)***

	Track 2—SSP						Track 2—Non-SSP						
	CPC+ mean[a]	C mean[a]	Impact estimate[b] (SE)	Percentage impact[c]	90 percent confidence interval	*p*-Value	CPC+ mean[a]	C mean[a]	Impact estimate[b] (SE)	Percentage impact[c]	90 percent confidence interval	*p*-Value	*p*-Value for SSP vs. non-SSP difference
PY 4	$23	$26	-$1.3*** ($0.5)	-5.1%	(-$2.0, -$0.5)	0.01	$25	$25	$0.1 ($0.3)	0.5%	(-$0.4, $0.7)	0.73	0.02
PY 5	$27	$30	-$1.6*** ($0.6)	-5.7%	(-$2.6, -$0.7)	0.01	$29	$30	-$0.2 ($0.4)	-0.8%	(-$1.0, $0.5)	0.58	0.06
PY 1 through 5	$26	$29	-$0.9** ($0.4)	-3.1%	(-$1.4, -$0.3)	0.02	$28	$28	$0.0 ($0.3)	-0.1%	(-$0.5, $0.4)	0.96	0.06
Expenditures for physician and nonphysician Part B noninstitutional services in any setting[j]													
Baseline	$248	$250	NA	NA	NA	NA	$243	$230	NA	NA	NA	NA	NA
PY 1	$251	$256	-$2.1* ($1.1)	-0.8%	(-$4.0, -$0.3)	0.05	$250	$235	$1.5 ($1.0)	0.6%	(-$0.2, $3.2)	0.16	0.02
PY 2	$265	$271	-$3.5* ($1.8)	-1.3%	(-$6.4, -$0.5)	0.05	$265	$249	$2.8** ($1.4)	1.1%	($0.6, $5.1)	0.04	0.01
PY 3	$278	$283	-$3.0 ($2.3)	-1.1%	(-$6.8, $0.8)	0.19	$279	$262	$3.7** ($1.7)	1.3%	($0.9, $6.5)	0.03	0.02
PY 4	$259	$262	-$1.3 ($2.4)	-0.5%	(-$5.3, $2.6)	0.58	$264	$245	$6.0*** ($2.0)	2.3%	($2.7, $9.3)	0.00	0.02
PY 5	$300	$306	-$3.7 ($2.7)	-1.2%	(-$8.1, $0.7)	0.17	$303	$282	$7.4*** ($2.4)	2.5%	($3.5, $11.3)	0.00	0.00
PY 1 through 5	$271	$276	-$2.8 ($1.8)	-1.0%	(-$5.7, $0.2)	0.13	$273	$256	$4.2*** ($1.4)	1.6%	($1.9, $6.6)	0.00	0.00
Expenditures for ambulatory visits with primary care practitioners													
Baseline	$24	$25	NA	NA	NA	NA	$24	$24	NA	NA	NA	NA	NA
PY 1	$25	$26	$0.2* ($0.1)	0.9%	($0.0, $0.5)	0.09	$25	$25	$0.0 ($0.1)	0.0%	(-$0.2, $0.2)	0.95	0.22
PY 2	$27	$26	$1.1*** ($0.2)	4.3%	($0.8, $1.4)	0.00	$27	$26	$0.9*** ($0.2)	3.6%	($0.6, $1.2)	0.00	0.55
PY 3	$28	$28	$1.3*** ($0.3)	4.8%	($0.9, $1.7)	0.00	$28	$27	$1.2*** ($0.2)	4.3%	($0.8, $1.6)	0.00	0.73
PY 4	$26	$25	$2.4*** ($0.3)	10.3%	($2.0, $2.9)	0.00	$26	$24	$2.0*** ($0.2)	8.2%	($1.6, $2.4)	0.00	0.26
PY 5	$32	$32	$1.4*** ($0.4)	4.5%	($0.8, $2.0)	0.00	$33	$31	$1.7*** ($0.3)	5.5%	($1.2, $2.2)	0.00	0.47
PY 1 through 5	$28	$27	$1.3*** ($0.2)	4.9%	($0.9, $1.6)	0.00	$28	$27	$1.2*** ($0.2)	4.3%	($0.8, $1.5)	0.00	0.66
Expenditures for ambulatory visits with primary care practitioners at assigned practice[k]													
Baseline	$17	$17	NA	NA	NA	NA	$17	$16	NA	NA	NA	NA	NA
PY 1	$17	$17	$0.4*** ($0.1)	2.5%	($0.2, $0.7)	0.00	$17	$16	$0.2 ($0.1)	1.3%	($0.0, $0.4)	0.11	0.32
PY 2	$17	$15	$1.7*** ($0.2)	11.4%	($1.4, $2.1)	0.00	$17	$14	$1.4*** ($0.2)	9.2%	($1.1, $1.7)	0.00	0.31
PY 3	$17	$15	$2.1*** ($0.3)	14.2%	($1.7, $2.6)	0.00	$17	$15	$1.9*** ($0.2)	12.0%	($1.5, $2.3)	0.00	0.44

Table 5.A.2.2b. ***(continued)***

	Track 2—SSP						Track 2—Non-SSP						
	CPC+ mean[a]	C mean[a]	Impact estimate[b] (SE)	Percentage impact[c]	90 percent confidence interval	*p*-Value	CPC+ mean[a]	C mean[a]	Impact estimate[b] (SE)	Percentage impact[c]	90 percent confidence interval	*p*-Value	*p*-Value for SSP vs. non-SSP difference
PY 4	$15	$12	$3.0*** ($0.3)	25.5%	($2.5, $3.5)	0.00	$15	$12	$2.6*** ($0.3)	20.7%	($2.1, $3.1)	0.00	0.43
PY 5	$17	$15	$2.5*** ($0.4)	17.5%	($1.9, $3.2)	0.00	$18	$14	$2.8*** ($0.4)	18.9%	($2.2, $3.4)	0.00	0.59
PY 1 through 5	$17	$15	$2.0*** ($0.2)	13.4%	($1.6, $2.4)	0.00	$17	$14	$1.8*** ($0.2)	11.7%	($1.4, $2.1)	0.00	0.56
Expenditures for ambulatory visits with primary care practitioners at non-assigned practice[k]													
Baseline	$7	$8	NA	NA	NA	NA	$7	$8	NA	NA	NA	NA	NA
PY 1	$8	$8	-$0.2* ($0.1)	-2.4%	(-$0.3, $0.0)	0.05	$8	$9	-$0.2*** ($0.1)	-2.9%	(-$0.4, -$0.1)	0.01	0.77
PY 2	$10	$11	-$0.6*** ($0.2)	-6.2%	(-$0.9, -$0.4)	0.00	$10	$11	-$0.5*** ($0.1)	-4.9%	(-$0.7, -$0.3)	0.00	0.50
PY 3	$11	$12	-$0.9*** ($0.2)	-7.3%	(-$1.2, -$0.5)	0.00	$11	$12	-$0.7*** ($0.2)	-6.0%	(-$1.0, -$0.4)	0.00	0.51
PY 4	$11	$12	-$0.6** ($0.2)	-4.7%	(-$1.0, -$0.1)	0.03	$11	$13	-$0.6*** ($0.2)	-5.2%	(-$1.0, -$0.2)	0.01	0.86
PY 5	$15	$17	-$1.1*** ($0.3)	-6.8%	(-$1.6, -$0.6)	0.00	$15	$17	-$1.1*** ($0.3)	-6.6%	(-$1.5, -$0.6)	0.00	0.89
PY 1 through 5	$11	$12	-$0.7*** ($0.2)	-5.7%	(-$1.0, -$0.4)	0.00	$11	$13	-$0.6*** ($0.1)	-5.3%	(-$0.8, -$0.4)	0.00	0.78
Expenditures for ambulatory visits with specialists													
Baseline	$26	$25	NA	NA	NA	NA	$23	$22	NA	NA	NA	NA	NA
PY 1	$25	$25	-$0.1 ($0.1)	-0.4%	(-$0.3, $0.1)	0.34	$23	$22	$0.0 ($0.1)	0.0%	(-$0.1, $0.1)	0.98	0.43
PY 2	$25	$25	-$0.4** ($0.1)	-1.4%	(-$0.6, -$0.1)	0.01	$23	$22	$0.0 ($0.1)	0.2%	(-$0.1, $0.2)	0.68	0.03
PY 3	$25	$26	-$0.5*** ($0.2)	-2.0%	(-$0.8, -$0.3)	0.00	$24	$23	$0.0 ($0.1)	0.1%	(-$0.2, $0.2)	0.93	0.01
PY 4	$22	$22	-$0.5*** ($0.2)	-2.4%	(-$0.8, -$0.2)	0.00	$20	$19	$0.0 ($0.1)	0.1%	(-$0.2, $0.3)	0.92	0.02
PY 5	$29	$29	-$0.7*** ($0.2)	-2.4%	(-$1.1, -$0.4)	0.00	$27	$26	$0.0 ($0.2)	0.0%	(-$0.3, $0.3)	0.96	0.02
PY 1 through 5	$25	$25	-$0.4*** ($0.1)	-1.7%	(-$0.7, -$0.2)	0.00	$24	$23	$0.0 ($0.1)	0.1%	(-$0.2, $0.2)	0.90	0.01
Expenditures on laboratory services[l]													
Baseline	$23	$24	NA	NA	NA	NA	$23	$23	NA	NA	NA	NA	NA
PY 1	$23	$24	-$0.5*** ($0.2)	-2.1%	(-$0.8, -$0.2)	0.00	$24	$24	-$0.1 ($0.1)	-0.4%	(-$0.3, $0.2)	0.54	0.06
PY 2	$25	$26	-$0.4 ($0.2)	-1.4%	(-$0.7, $0.0)	0.10	$26	$26	-$0.3 ($0.3)	-1.3%	(-$0.8, $0.1)	0.23	0.94
PY 3	$26	$26	-$0.1 ($0.3)	-0.3%	(-$0.5, $0.3)	0.74	$26	$26	$0.0 ($0.2)	-0.1%	(-$0.4, $0.4)	0.91	0.86

Table 5.A.2.2b. ***(continued)***

	Track 2—SSP						Track 2—Non-SSP						
	CPC+ mean[a]	C mean[a]	Impact estimate[b] (SE)	Percentage impact[c]	90 percent confidence interval	*p*-Value	CPC+ mean[a]	C mean[a]	Impact estimate[b] (SE)	Percentage impact[c]	90 percent confidence interval	*p*-Value	*p*-Value for SSP vs. non-SSP difference
PY 4	$28	$28	-$0.2 ($0.3)	-0.6%	(-$0.7, $0.4)	0.62	$28	$28	-$0.4 ($0.2)	-1.2%	(-$0.8, $0.1)	0.16	0.62
PY 5	$32	$34	-$1.6*** ($0.4)	-4.6%	(-$2.2, -$1.0)	0.00	$32	$33	-$0.8*** ($0.3)	-2.4%	(-$1.3, -$0.3)	0.01	0.11
PY 1 through 5	$27	$28	-$0.5** ($0.2)	-1.9%	(-$0.9, -$0.2)	0.02	$27	$27	-$0.3* ($0.2)	-1.1%	(-$0.6, $0.0)	0.09	0.47
Expenditures on imaging services[l]													
Baseline	$46	$46	NA	NA	NA	NA	$44	$44	NA	NA	NA	NA	NA
PY 1	$47	$48	-$0.3 ($0.4)	-0.6%	(-$0.9, $0.3)	0.43	$45	$45	$0.0 ($0.3)	0.1%	(-$0.5, $0.6)	0.94	0.51
PY 2	$49	$51	-$0.8* ($0.4)	-1.6%	(-$1.5, -$0.1)	0.05	$48	$48	$0.3 ($0.3)	0.6%	(-$0.3, $0.8)	0.40	0.04
PY 3	$53	$54	-$0.8* ($0.5)	-1.6%	(-$1.6, $0.0)	0.09	$51	$51	$0.2 ($0.4)	0.4%	(-$0.4, $0.8)	0.62	0.09
PY 4	$46	$48	-$1.2** ($0.6)	-2.4%	(-$2.1, -$0.2)	0.04	$46	$46	$0.0 ($0.4)	-0.1%	(-$0.7, $0.6)	0.93	0.10
PY 5	$53	$55	-$1.6*** ($0.6)	-3.0%	(-$2.6, -$0.7)	0.00	$52	$52	-$0.2 ($0.5)	-0.5%	(-$1.0, $0.5)	0.61	0.06
PY 1 through 5	$50	$51	-$0.9** ($0.4)	-1.9%	(-$1.6, -$0.3)	0.02	$49	$49	$0.1 ($0.3)	0.1%	(-$0.4, $0.5)	0.85	0.04
Skilled nursing facility expenditures													
Baseline	$69	$69	NA	NA	NA	NA	$62	$60	NA	NA	NA	NA	NA
PY 1	$68	$66	$1.0 ($1.1)	1.5%	(-$0.7, $2.8)	0.33	$60	$59	-$0.6 ($1.0)	-1.0%	(-$2.3, $1.1)	0.55	0.27
PY 2	$68	$66	$1.4 ($1.1)	2.1%	(-$0.4, $3.3)	0.21	$61	$60	-$0.1 ($1.0)	-0.2%	(-$1.8, $1.5)	0.89	0.31
PY 3	$67	$66	$0.9 ($1.3)	1.4%	(-$1.2, $3.0)	0.47	$60	$59	$0.0 ($1.1)	0.0%	(-$1.9, $1.9)	0.98	0.60
PY 4	$66	$66	-$0.9 ($1.5)	-1.4%	(-$3.3, $1.5)	0.53	$61	$59	$1.4 ($1.2)	2.3%	(-$0.6, $3.4)	0.26	0.22
PY 5	$68	$66	$1.6 ($1.6)	2.4%	(-$1.1, $4.3)	0.32	$63	$59	$2.6** ($1.3)	4.4%	($0.5, $4.8)	0.04	0.62
PY 1 through 5	$67	$66	$0.8 ($1.1)	1.2%	(-$1.0, $2.6)	0.45	$61	$59	$0.6 ($0.9)	1.0%	(-$0.9, $2.1)	0.50	0.89
Home health expenditures													
Baseline	$41	$44	NA	NA	NA	NA	$41	$40	NA	NA	NA	NA	NA
PY 1	$40	$43	-$0.1 ($0.5)	-0.2%	(-$0.8, $0.7)	0.86	$41	$40	-$0.5 ($0.4)	-1.3%	(-$1.2, $0.2)	0.24	0.49
PY 2	$40	$45	-$0.9* ($0.5)	-2.2%	(-$1.7, -$0.1)	0.07	$41	$40	-$0.6 ($0.5)	-1.5%	(-$1.4, $0.2)	0.21	0.70
PY 3	$40	$44	-$0.4 ($0.6)	-0.9%	(-$1.3, $0.6)	0.52	$41	$41	-$1.4** ($0.5)	-3.2%	(-$2.2, -$0.5)	0.01	0.20

Table 5.A.2.2b. ***(continued)***

	Track 2—SSP						Track 2—Non-SSP						
	CPC+ mean[a]	C mean[a]	Impact estimate[b] (SE)	Percentage impact[c]	90 percent confidence interval	*p*-Value	CPC+ mean[a]	C mean[a]	Impact estimate[b] (SE)	Percentage impact[c]	90 percent confidence interval	*p*-Value	*p*-Value for SSP vs. non-SSP difference
PY 4	$36	$41	-$1.6** ($0.7)	-4.1%	(-$2.6, -$0.5)	0.02	$38	$38	-$1.7*** ($0.6)	-4.3%	(-$2.7, -$0.7)	0.00	0.84
PY 5	$37	$43	-$2.8*** ($0.7)	-7.0%	(-$4.0, -$1.7)	0.00	$40	$40	-$1.5** ($0.6)	-3.7%	(-$2.5, -$0.5)	0.01	0.18
PY 1 through 5	$39	$43	-$1.1** ($0.5)	-2.8%	(-$1.9, -$0.3)	0.02	$40	$40	-$1.1** ($0.4)	-2.8%	(-$1.9, -$0.4)	0.01	0.96
Hospice expenditures													
Baseline	$22	$23	NA	NA	NA	NA	$25	$27	NA	NA	NA	NA	NA
PY 1	$23	$24	$0.4 ($0.6)	1.6%	(-$0.6, $1.4)	0.55	$25	$26	$0.9 ($0.6)	3.8%	($0.0, $1.9)	0.10	0.50
PY 2	$26	$26	$2.0** ($0.8)	8.0%	($0.7, $3.3)	0.01	$29	$28	$2.8*** ($0.7)	10.5%	($1.6, $4.0)	0.00	0.46
PY 3	$30	$28	$3.2*** ($0.9)	11.9%	($1.7, $4.7)	0.00	$32	$31	$3.2*** ($0.8)	10.9%	($1.9, $4.5)	0.00	0.98
PY 4	$32	$31	$2.2** ($0.9)	7.4%	($0.7, $3.7)	0.01	$34	$33	$2.1** ($0.8)	6.5%	($0.7, $3.4)	0.01	0.91
PY 5	$33	$32	$2.6*** ($1.0)	8.5%	($1.0, $4.2)	0.01	$35	$34	$3.0*** ($1.0)	9.1%	($1.4, $4.6)	0.00	0.77
PY 1 through 5	$29	$28	$2.1*** ($0.7)	7.8%	($0.9, $3.3)	0.00	$32	$31	$2.4*** ($0.6)	8.3%	($1.4, $3.5)	0.00	0.74
Durable medical equipment expenditures													
Baseline	$20	$20	NA	NA	NA	NA	$21	$22	NA	NA	NA	NA	NA
PY 1	$19	$19	$0.1 ($0.3)	0.4%	(-$0.5, $0.6)	0.83	$20	$20	$0.5 ($0.3)	2.4%	(-$0.1, $1.0)	0.16	0.40
PY 2	$22	$22	$0.0 ($0.5)	-0.1%	(-$0.8, $0.7)	0.97	$23	$23	$0.5 ($0.4)	2.2%	(-$0.1, $1.1)	0.18	0.38
PY 3	$24	$23	$0.2 ($0.5)	0.8%	(-$0.6, $1.0)	0.72	$25	$24	$0.8* ($0.4)	3.2%	($0.1, $1.4)	0.05	0.35
PY 4	$25	$24	$0.5 ($0.5)	2.2%	(-$0.3, $1.4)	0.30	$26	$26	$0.7 ($0.5)	2.6%	(-$0.1, $1.4)	0.16	0.86
PY 5	$26	$25	-$0.2 ($0.5)	-0.8%	(-$1.0, $0.6)	0.67	$26	$26	$0.7 ($0.5)	2.9%	($0.0, $1.5)	0.10	0.16
PY 1 through 5	$23	$23	$0.1 ($0.4)	0.5%	(-$0.5, $0.7)	0.78	$24	$24	$0.6* ($0.3)	2.7%	($0.1, $1.2)	0.07	0.32
Unweighted sample sizes[m]													
Number of practices	636	1,817					879	1,966					
Number of beneficiaries	847,208	2,257,322					1,053,634	2,261,852					
Number of beneficiary-years	3,204,963	8,538,135					4,020,326	8,516,384					

Source: Mathematica's analysis of Medicare claims data from January 2013 through December 2021.

Table 5.A.2.2b. ***(continued)***

Notes: Although this table indicates which estimates are statistically significant, when we interpret evidence, we combine evidence from the magnitude of the effect, the *p*-values, findings on related outcomes, subgroups, sensitivity tests, and other data sources about model implementation.

[a] We report the actual, unadjusted averages in the baseline period which are similar for the CPC+ and comparison groups due to matching. In the intervention periods, the comparison group mean is computed by subtracting the regression adjusted difference between the CPC+ and comparison means in each time period from the CPC+ mean in that same time period.

[b] Each impact estimate is regression-adjusted using a difference-in-differences analysis that reflects the difference of the average outcome for Medicare FFS beneficiaries attributed to CPC+ practices in the first five years of CPC+ to the average outcome in the baseline year, relative to the same difference over time for Medicare FFS beneficiaries attributed to comparison practices, while controlling for beneficiary characteristics, practice fixed effects and COVID-19 related controls.

[c] We calculated percentage impacts relative to what the CPC+ mean would have been in Program Years 1 through 5 (separately and combined) in the absence of the intervention—that is, the unadjusted CPC+ mean minus the impact estimate.

[d] Expenditures for Part A and Part B services in PY 3, PY 4, and PY 5 include QPP payment adjustments, based on practitioner performance two years before. They are applicable for both CPC+ and comparison practices. The adjustments are composed of (1) MIPS adjustments, which are applied directly to physician and outpatient claims (as a percentage of the charges on the claims); and (2) lump sum incentive payments to eligible practitioners who participated in Advanced APMs in 2017, 2018, and 2019 (calculated based on 2018, 2019, and 2020 claims for these practitioners, respectively). The first QPP adjustments were paid in PY 3 (two years after the start of QPP), so there are no QPP payments in PYs 1 and 2.

[e] We determine SSP ACO participation status based on participation at the beginning of PY 1 (January 1, 2017). However, over time, CPC+ practices may join or leave SSP, resulting in a small subset of SSP practices receiving the Performance-based Incentive Payments and a small subset of non-SSP practices receiving the shared savings payments. This is reflected in the impact estimates.

[f] The sum of expenditures by claim type does not equal the total expenditures for Part A and B services without enhanced payments in PY 3, PY 4, and PY 5 because the total expenditures include lum*p*-sum incentive payments that are not applied at the claim level and are instead paid out directly to eligible practitioners who participated in Advanced APMs in 2017, 2018 and 2019.

[g] Acute inpatient care includes short-stay acute hospital admissions and admissions to CAHs. Expenditures for non-acute hospital admissions other than those for inpatient rehabilitation, such as psychiatric hospital admissions, are included in inpatient expenditures but not shown separately.

[h] Post-acute care expenditures include expenditures on home health, long-term care, skilled nursing facility, and inpatient rehabilitation. These are not a sub-category of inpatient expenditures.

[i] Expenditures, with QPP payment adjustments, on outpatient ED visits include professional (which is part of expenditures for physician and nonphysician Part B noninstitutional services) and facility fees, as well as payments for observation stays.

[j] Expenditures, with QPP payment adjustments, on Part B noninstitutional services include expenditures for (1) ambulatory primary care visits, (2) ambulatory specialist visits, and (3) non-ambulatory physician visits as well as services provided by other noninstitutional providers. (We only show the first two categories separately in the table).

[k] We define the assigned practice for the baseline period as the first practice to which a beneficiary was attributed during the baseline period, and the assigned practice for the intervention period as the first practice that the beneficiary was attributed to during the intervention period.

[l] Laboratory and imaging services were identified in both the carrier and outpatient claim files.

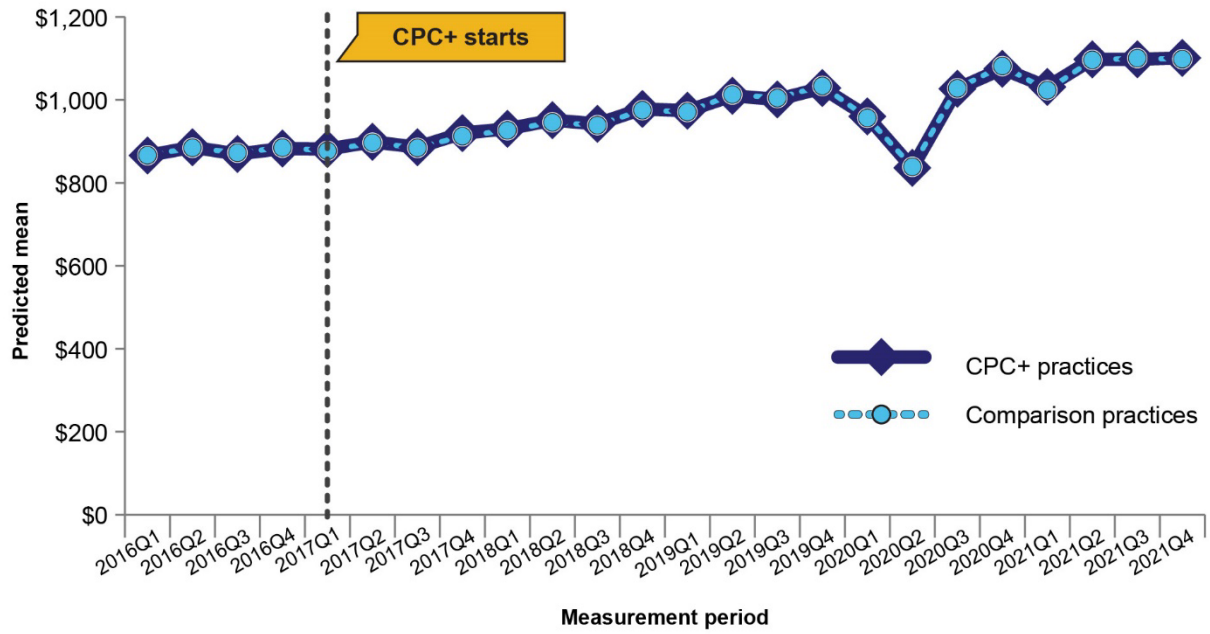
[m] After accounting for weights that adjust for matching and time observed in Medicare FFS, the effective sample sizes fall but are still substantial. For the comparison group, the effective sample size is 38 to 43 percent of the actual sample size. The effective sample size for the CPC+ group is about 96 percent of the actual sample size because it is affected only by time observed (and not by the matching weights).

*/**/*** Significantly different from zero at the 0.10/0.05/0.01 level, two-tailed test.

NA = not applicable because the difference-in-differences impact estimate cannot be calculated at baseline.

ACO = Accountable Care Organization; APM = Alternative Payment Model; C = comparison; CAH = critical access hospital; ED = emergency department; FFS = fee-for-service; MIPS = Merit-based Incentive Payment System; NA = not applicable; PY = Program Year; QPP = Quality Payment Program; SE = standard error; SSP = Medicare Shared Savings Program.

Figure 5.A.2.3. Quarterly trends in average Medicare expenditures PBPM, excluding CMS's enhanced payments, Track 2

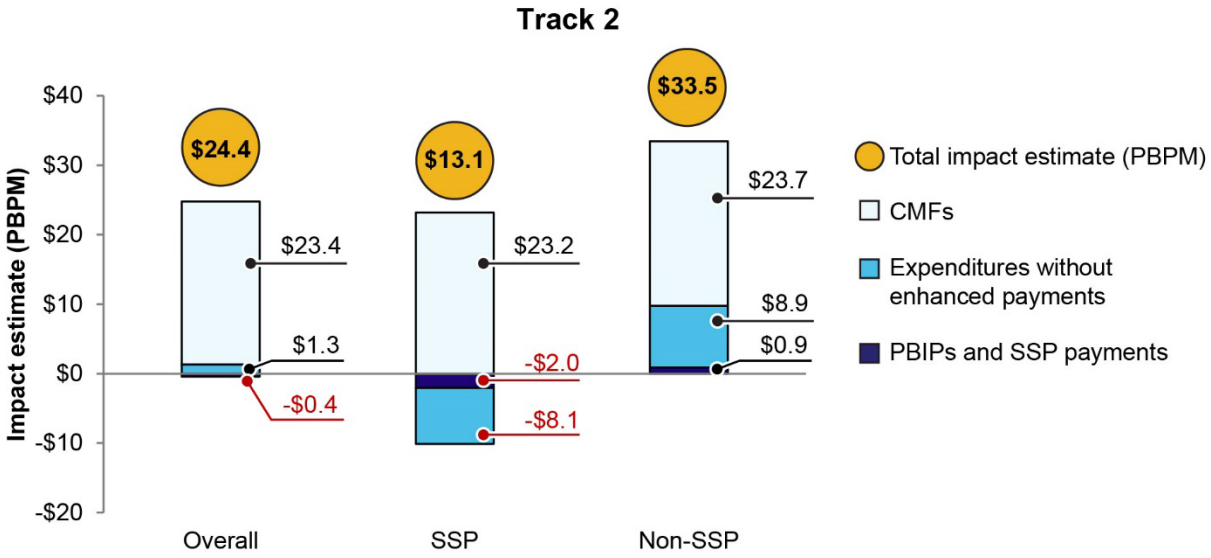


Source: Analyses of Medicare claims data from January 2013 through December 2021.

Notes: For beneficiaries attributed to CPC+ and comparison practices, the figure shows actual, unadjusted average expenditures in the baseline quarters (Q1 through Q4 of 2016), which are similar for the two groups due to matching. In the intervention quarters (starting in Q1 2017), the comparison group mean is regression-adjusted based on the quarterly difference-in-differences model, which controls for baseline characteristics and COVID-19 related controls. The sharp decline in expenditures during the first and second quarters of 2020 can be attributed to a decline in the overall utilization of health services during the initial months of the COVID-19 pandemic.

PBPM = per beneficiary per month.

Figure 5.A.2.4. Per beneficiary per month impact estimates for Medicare expenditures, with CMS's enhanced payments, by Track 2 and SSP status



Source: Mathematica's analysis of Medicare claims data from January 2013 through December 2021.

Notes: The impact estimates on expenditures without enhanced payments over the five years of CPC+ ($1.3 in Track 2 overall, -$8.1 in Track 2 SSP, $8.9 in Track 2 non-SSP) were not statistically significant overall or for the SSP subgroup but were statistically significant for the non-SSP subgroup. The impact estimates on expenditures including enhanced payments that are attributable to PBIPs and SSP payments were smaller by $0.4 for Track 2 overall and by $2.0 for Track 2 SSP, compared to the respective impact estimates that do not include PBIPs and SSP payments. The estimates attributable to PBIPs and SSP payments are negative because, between the baseline and the intervention period, the change due to PBIPs were $26.1 and $15.6 higher for CPC+ practices than for comparison practices in Track 2 overall and Track 2 SSP, respectively (because only CPC+ practices receive PBIPs) and the change due to SSP payments were $26.6 and $17.7 lower for CPC+ practices than for comparison practices, for Track 2 overall and Track 2 SSP respectively. This resulted in the impact estimates decreasing by $0.4 and $2.0, for Track 2 overall and Track 2 SSP respectively, after including both PBIPs and SSP payments.

Figure includes: (1) 636 Track 2 SSP and 879 Track 2 non-SSP CPC+ practices that were participating in CPC+ as of April 1, 2017 (the end of the first program quarter), and (2) 1,817 Track 2 SSP and 1,966 Track 2 non-SSP comparison practices.

CMF = care management fee; CMS = Centers for Medicare and Medicaid Services; PBIP = Performance-based Incentive Payment; PBPM = per beneficiary per month; SSP = Medicare Shared Savings Program.

Table 5.A.2.3a. Estimated average annual impacts of CPC+ on Medicare expenditures without CMS's enhanced payments across the five program years, by baseline practice characteristics, Track 1

Practice subgroup definition, based on baseline characteristics	Track 1 – Overall: Number (percentage) of CPC+ beneficiaries in subgroup at baseline	Track 1 – Overall: Impact estimate (standard error)	Track 1 – Overall: Percentage impact[a]	Track 1 – Overall: *p*-Value for difference in impact estimates between subgroups[b]
Main analysis (all practices)- average annual estimate for PY 1 through PY 5	**-**	**$1.1 ($3.3)**	**0.1%**	**-**
Whether practice participated in prior primary care transformation initiatives (recognized as a medical home or participated in MAPCP or CPC Classic)				
Yes	468,487 (53.6%)	$1.0 ($4.6)	0.1%	
No	405,383 (46.4%)	$0.8 ($4.7)	0.1%	0.92
Large and medium versus small practice based on number of primary care practitioners				
Large (6+ primary care practitioners)	404,456 (46.3%)	-$4.5 ($5.0)	-0.5%	
Medium (3–5 primary care practitioners)	282,380 (32.3%)	$9.9 ($5.7)	1.0%	
Small (1–2 primary care practitioners)	187,034 (21.4%)	-$0.7 ($7.1)	-0.1%	0.12
Whether hospital- or system-owned versus independent (based on IQVIA data)[c]				
Hospital- or system-owned	474,606 (54.3%)	$4.5 ($4.5)	0.5%	
Independent	399,264 (45.7%)	-$3.4 ($4.8)	-0.3%	0.26
Whether the practice shared a TIN with another primary care practice[c]				
Shared a TIN with another primary care practice	684,507 (78.3%)	$2.3 ($3.8)	0.2%	
Did not share a TIN with another primary care practice	189,364 (21.7%)	-$2.6 ($6.8)	-0.3%	0.55
Practice type: multi-specialty versus primary care only				
Multi-specialty	170,691 (19.5%)	$6.6 ($8.6)	0.7%	
Primary care only	703,179 (80.5%)	-$0.4 ($3.5)	0.0%	0.22
Urbanicity of practice's county: rural or suburban location versus urban location				
Rural	89,834 (10.3%)	-$0.2 ($10.0)	0.0%	
Suburban	156,799 (17.9%)	$8.1 ($8.4)	0.9%	
Urban	627,237 (71.8%)	-$0.7 ($3.9)	-0.1%	0.69

Source: Mathematica's analysis of Medicare claims data from January 2013 through December 2021.

Note: The estimates (and standard errors) in the impact estimate column show subgroup-specific impacts over the five years of CPC+, separately, for each practice characteristic listed in the table. We only tested differences within each subgroup if the estimates were significantly different between the two subgroups (that is, the *p*-value in the last column was <.10). Asterisks denote whether the impact estimate within a subgroup was significantly different from zero when estimates were significantly different between the subgroup categories.

[a] We calculated percentage impacts relative to what the CPC+ mean would have been in Program Years 1 through 5 (separately and combined) in the absence of the intervention—that is, the unadjusted CPC+ mean minus the impact estimate.

[b] The *p*-values in the last column represent results from testing for statistically significant differences in impact estimates between the subgroups, based on the baseline practice characteristic (using a t-test for subgroups with two categories and from an F-test for subgroups with more than two categories).

[c] Since ownership status of a practice is likely to be highly correlated with whether the practice shares TIN with other practices, we included only one of these characteristics at a time in these regressions.

*/**/***Within-subgroup estimate significantly different from zero at the 0.10/0.05/0.01 level, two-tailed test.

CMS = Centers for Medicare and Medicaid Services; CPC = Comprehensive Primary Care; MAPCP = Multi-payer Advanced Primary Care Practice Demonstration; PY = Program Year; TIN = Tax Identification Number

Table 5.A.2.3b. Estimated average annual impacts of CPC+ on Medicare expenditures without CMS's enhanced payments across the five program years, by baseline practice characteristics and SSP status, Track 1

Practice subgroup definition, based on baseline characteristics	Track 1 – SSP				Track 1 – Non-SSP			
	Number (percentage) of CPC+ beneficiaries in subgroup at baseline	Impact estimate (standard error)	Percentage impact[a]	*p*-Value for difference in impact estimates between subgroups[a]	Number (percentage) of CPC+ beneficiaries in subgroup at baseline	Impact estimate (standard error)	Percentage impact[a]	*p*-Value for difference in impact estimates between subgroups[b]
Main analysis (all practices)- average annual estimate for PY 1 through PY 5	**-**	**-$7.8* ($4.5)**	**-0.8%**	**-**	**-**	**$10.1** ($4.8)**	**1.1%**	**-**
Whether practice participated in prior primary care transformation initiatives (recognized as a medical home or participated in MAPCP or CPC Classic)								
Yes	214,075 (47.7%)	-$1.9 ($6.4)	-0.2%		254,262 (59.8%)	$3.4 ($6.5)	0.4%	
No	234,948 (52.3%)	-$13.3** ($6.2)	-1.3%	0.08	170,586 (40.2%)	$18.2 ($7.1)	1.9%	0.12
Large and medium versus small practice based on number of primary care practitioners								
Large (6+ primary care practitioners)	189,229 (42.1%)	-$11.4 ($7.2)	-1.1%		215,122 (50.6%)	-$0.7 ($6.9)	-0.1%	
Medium (3–5 primary care practitioners)	156,338 (34.8%)	-$3.5 ($6.9)	-0.4%		126,106 (29.7%)	$27.2 ($9.3)	2.9%	
Small (1–2 primary care practitioners)	103,455 (23.0%)	-$8.0 ($9.8)	-0.8%	0.71	83,621 (19.7%)	$8.3 ($10.1)	0.9%	0.10
Whether hospital- or system-owned versus independent (based on IQVIA data)[c]								
Hospital- or system-owned	250,558 (55.8%)	-$7.2 ($5.9)	-0.7%		224,086 (52.7%)	$16.5** ($7.0)	1.7%	
Independent	198,464 (44.2%)	-$8.8 ($6.9)	-0.9%	0.54	200,762 (47.3%)	$1.4 ($6.6)	0.2%	0.08
Whether the practice shared a TIN with another primary care practice[c]								
Shared a TIN with another primary care practice	366,843 (81.7%)	-$5.7 ($5.0)	-0.6%		317,749 (74.8%)	$10.8 ($5.7)	1.2%	
Did not share a TIN with another primary care practice	82,179 (18.3%)	-$17.3 ($10.2)	-1.7%	0.39	107,099 (25.2%)	$8.1 ($8.8)	0.9%	0.53
Practice type: multi-specialty versus primary care only								
Multi-specialty	76,547 (17.0%)	-$1.3 ($10.9)	-0.1%		94,082 (22.1%)	$10.7 ($12.5)	1.2%	
Primary care only	372,475 (83.0%)	-$9.2 ($4.9)	-0.9%	0.36	330,766 (77.9%)	$9.0 ($5.1)	0.9%	0.48
Urbanicity of practice's county: rural or suburban location versus urban location								
Rural	22,327 (5.0%)	-$28.0 ($20.7)	-3.0%		67,372 (15.9%)	$9.8 ($11.7)	1.1%	
Suburban	74,982 (16.7%)	$15.3 ($10.0)	1.6%		81,785 (19.3%)	$0.5 ($13.1)	0.1%	
Urban	351,712 (78.3%)	-$11.5 ($5.2)	-1.1%	0.01	275,691 (64.9%)	$11.9 ($5.9)	1.2%	0.69

Source: Mathematica's analysis of Medicare claims data from January 2013 through December 2021.

Note: The estimates (and standard errors) in the impact estimate column show subgrou*p*-specific impacts over the five years of CPC+, separately, for each practice characteristic listed in the table. We only tested differences within each subgroup if the estimates were significantly different between the two subgroups (that is, the *p*-value in the last

Table 5.A.2.3b. ***(continued)***

column was <.10). Asterisks denote whether the impact estimate within a subgroup was significantly different from zero when estimates were significantly different between the subgroup categories.

[a] We calculated percentage impacts relative to what the CPC+ mean would have been in Program Years 1 through 5 (separately and combined) in the absence of the intervention—that is, the unadjusted CPC+ mean minus the impact estimate.

[b] The *p*-values in the last column represent results from testing for statistically significant differences in impact estimates between the subgroups, based on the baseline practice characteristic (using a t-test for subgroups with two categories and from an F-test for subgroups with more than two categories).

[c] Since ownership status of a practice is likely to be highly correlated with whether the practice shares TIN with other practices, we included only one of these characteristics at a time in these regressions.

*/**/***Within-subgroup estimate significantly different from zero at the 0.10/0.05/0.01 level, two-tailed test.

CMS = Centers for Medicare and Medicaid Services; CPC = Comprehensive Primary Care; MAPCP = Multi-payer Advanced Primary Care Practice Demonstration; PY = Program Year; SSP = Medicare Shared Savings Program; TIN = Tax Identification Number

Table 5.A.2.4a. Estimated average annual impacts of CPC+ on Medicare expenditures without CMS's enhanced payments across the five program years, by baseline practice characteristics, Track 2

Practice subgroup definition, based on baseline characteristics	Track 2 – Overall			
	Number (percentage) of CPC+ beneficiaries in subgroup at baseline	Impact estimate (standard error)	Percentage impact	*p*-Value for difference in impact estimates between subgroups[a]
Main analysis (all practices)- average annual estimate for PY 1 through PY 5	-	**$1.3 ($3.8)**	**0.1%**	
Whether practice participated in prior primary care transformation initiatives (recognized as a medical home or participated in MAPCP or CPC Classic)				
Yes	865,798 (81.2%)	-$3.2 ($4.4)	-0.3%	
No	201,028 (18.8%)	$18.9** ($7.5)	1.9%	0.02
Large and medium versus small practice based on number of primary care practitioners				
Large (6+ primary care practitioners)	589,224 (55.2%)	-$5.1 ($5.5)	-0.5%	
Medium (3–5 primary care practitioners)	340,406 (31.9%)	$8.7 ($6.1)	0.9%	
Small (1–2 primary care practitioners)	137,196 (12.9%)	$7.7 ($9.2)	0.8%	0.33
Whether hospital- or system-owned versus independent (based on IQVIA data)[b]				
Hospital- or system-owned	619,957 (58.1%)	$6.3 ($5.0)	0.6%	
Independent	446,869 (41.9%)	-$6.4 ($5.6)	-0.7%	0.05
Whether the practice shared a TIN with another primary care practice[b]				
Shared a TIN with another primary care practice	913,196 (85.6%)	$3.7 ($4.2)	0.4%	
Did not share a TIN with another primary care practice	153,630 (14.4%)	-$10.5 ($8.8)	-1.1%	0.03
Practice type: multi-specialty versus primary care only				
Multi-specialty	278,801 (26.1%)	-$2.3 ($9.1)	-0.2%	
Primary care only	788,025 (73.9%)	$2.1 ($4.0)	0.2%	0.98
Urbanicity of practice's county: rural or suburban location versus urban location				
Rural	82,613 (7.7%)	-$0.3 ($11.5)	0.0%	
Suburban	170,323 (16.0%)	$4.9 ($12.0)	0.5%	
Urban	813,890 (76.3%)	$0.2 ($4.1)	0.0%	0.93

Source: Mathematica's analysis of Medicare claims data from January 2013 through December 2021.

Note: The estimates (and standard errors) in the impact estimate column show subgrou*p*-specific impacts over the five years of CPC+, separately, for each practice characteristic listed in the table. We only tested differences within each subgroup if the estimates were significantly different between the two subgroups (that is, the *p*-value in the last column was <.10). Asterisks denote whether the impact estimate within a subgroup was significantly different from zero when estimates were significantly different between the subgroup categories.

[a] The *p*-values in the last column represent results from testing for statistically significant differences in impact estimates between the subgroups, based on the baseline practice characteristic (using a t-test for subgroups with two categories and from an F-test for subgroups with more than two categories).

[b] Since ownership status of a practice is likely to be highly correlated with whether the practice shares TIN with other practices, we included only one of these characteristics at a time in these regressions.

*/**/***Within-subgroup estimate significantly different from zero at the 0.10/0.05/0.01 level, two-tailed test.

CMS = Centers for Medicare and Medicaid Services; CPC = Comprehensive Primary Care; MAPCP = Multi-payer Advanced Primary Care Practice Demonstration; PY = Program Year; TIN = Tax Identification Number

Table 5.A.2.4b. Estimated average annual impacts of CPC+ on Medicare expenditures without CMS's enhanced payments across the five program years, by baseline practice characteristics and SSP status, Track 2

Practice subgroup definition, based on baseline characteristics	Track 2 – SSP				Track 2– Non-SSP			
	Number (percentage) of CPC+ beneficiaries in subgroup at baseline	Impact estimate (standard error)	Percentage impact	*p*-Value for difference in impact estimates between subgroups[a]	Number (percentage) of CPC+ beneficiaries in subgroup at baseline	Impact estimate (standard error)	Percentage impact	*p*-Value for difference in impact estimates between subgroups[a]
Main analysis (all practices)- average annual estimate for PY 1 through PY 5	**-**	**-$8.1 ($6.1)**	**-0.8%**	**-**	**-**	**$8.9* ($4.7)**	**0.9%**	**-**
Whether practice participated in prior primary care transformation initiatives (recognized as a medical home or participated in MAPCP or CPC Classic)								
Yes	385,875 (81.8%)	-$15.1** ($6.5)	-1.5%		479,947 (80.6%)	$7.4 ($5.6)	0.8%	
No	85,762 (18.2%)	$24.1* ($13.6)	2.3%	0.02	115,242 (19.4%)	$14.5 ($8.3)	1.5%	0.52
Large and medium versus small practice based on number of primary care practitioners								
Large (6+ primary care practitioners)	279,067 (59.2%)	-$11.7 ($8.3)	-1.2%		310,301 (52.1%)	$2.4 ($6.7)	0.3%	
Medium (3–5 primary care practitioners)	134,103 (28.4%)	-$3.1 ($9.2)	-0.3%		206,177 (34.6%)	$16.6 ($8.1)	1.8%	
Small (1–2 primary care practitioners)	58,467 (12.4%)	-$1.3 ($13.3)	-0.1%	0.84	78,712 (13.2%)	$13.4 ($12.6)	1.4%	0.26
Whether hospital- or system-owned versus independent (based on IQVIA data)[b]								
Hospital- or system-owned	289,350 (61.4%)	-$8.1 ($7.4)	-0.8%		330,724 (55.6%)	$18.9*** ($6.5)	2.0%	
Independent	182,287 (38.6%)	-$7.8 ($9.3)	-0.8%	0.86	264,465 (44.4%)	-$3.9 ($6.9)	-0.4%	0.01
Whether the practice shared a TIN with another primary care practice[b]								
Shared a TIN with another primary care practice	416,348 (88.3%)	-$5.0 ($6.2)	-0.5%		496,945 (83.5%)	$12.3 ($5.4)	1.3%	
Did not share a TIN with another primary care practice	55,289 (11.7%)	-$23.0 ($15.0)	-2.3%	0.09	98,244 (16.5%)	-$4.7 ($10.7)	-0.5%	0.10
Practice type: multi-specialty versus primary care only								
Multi-specialty	116,601 (24.7%)	-$13.2 ($14.2)	-1.3%		162,149 (27.2%)	$9.0 ($11.5)	0.9%	
Primary care only	355,036 (75.3%)	-$6.2 ($6.3)	-0.6%	0.77	433,040 (72.8%)	$8.7 ($5.3)	0.9%	0.60
Urbanicity of practice's county: rural or suburban location versus urban location								
Rural	18,533 (3.9%)	-$19.6 ($19.5)	-2.2%		63,941 (10.7%)	$6.1 ($13.6)	0.7%	
Suburban	75,938 (16.1%)	-$2.9 ($19.5)	-0.3%		94,390 (15.9%)	$15.1 ($14.2)	1.7%	
Urban	377,166 (80.0%)	-$8.4 ($6.0)	-0.8%	0.83	436,858 (73.4%)	$7.8 ($5.5)	0.8%	0.88

Source: Mathematica's analysis of Medicare claims data from January 2013 through December 2021.

Note: The estimates (and standard errors) in the impact estimate column show subgroup-specific impacts over the five years of CPC+, separately, for each practice characteristic listed in the table. We only tested differences within each subgroup if the estimates were significantly different between the two subgroups (that is, the *p*-value in the last

Table 5.A.2.4b. ***(continued)***

column was <.10). Asterisks denote whether the impact estimate within a subgroup was significantly different from zero when estimates were significantly different between the subgroup categories.

[a] The *p*-values in the last column represent results from testing for statistically significant differences in impact estimates between the subgroups, based on the baseline practice characteristic (using a t-test for subgroups with two categories and from an F-test for subgroups with more than two categories).

[b] Since ownership status of a practice is likely to be highly correlated with whether the practice shares TIN with other practices, we included only one of these characteristics at a time in these regressions.

*/**/***Within-subgroup estimate significantly different from zero at the 0.10/0.05/0.01 level, two-tailed test.

CMS = Centers for Medicare and Medicaid Services; CPC = Comprehensive Primary Care; MAPCP = Multi-payer Advanced Primary Care Practice Demonstration; PY = Program Year; SSP = Medicare Shared Savings Program; TIN = Tax Identification Number

Table 5.A.2.5b. ***(continued)***

Table 5.A.2.5a. Estimated average annual impacts of CPC+ on Medicare expenditures without CMS's enhanced payments across the five program years, by baseline beneficiary characteristics, Track 1

Beneficiary subgroup definition, based on baseline characteristics	Track 1 - Overall			
	Number (percentage) of CPC+ beneficiaries in subgroup at baseline	Impact estimate (standard error)	Percentage impact[a]	*p*-Value for difference in impact estimates between subgroups[b]
Main analysis (all beneficiaries)- average annual estimate for PY 1 through PY 5	**-**	**$1.1 ($3.3)**	**0.1%**	**-**
Patients in the highest quartile of the HCC score distribution				
Yes	203,811 (25.9%)	-$4.1 ($9.7)	-0.2%	
No	583,156 (74.1%)	$4.5 ($2.8)	0.6%	0.38
Patients in the highest decile of the HCC score distribution or who have dementia				
Yes	123,085 (15.6%)	$0.0 ($13.4)	0.0%	
No	663,882 (84.4%)	$2.2 ($3.0)	0.3%	0.87
Patients with anxiety/depression or substance use disorders				
Yes	120,562 (16.6%)	$3.0 ($9.4)	0.2%	
No	604,012 (83.4%)	$3.7 ($3.4)	0.4%	0.94
Patients with multiple chronic conditions (at least 2 of 12 frequently occurring chronic conditions) and one or more hospitalizations[d]				
Yes	68,204 (8.7%)	$3.1 ($18.9)	0.1%	
No	718,763 (91.3%)	$2.0 ($3.1)	0.2%	0.95
Patients dually eligible for Medicare and Medicaid				
Yes	107,885 (12.6%)	-$13.1 ($10.8)	-1.0%	
No	746,776 (87.4%)	$3.1 ($3.3)	0.3%	0.14

Source: Mathematica's analysis of Medicare claims data from January 2013 through December 2021.

Note: Beneficiary characteristics to determine subgroup membership are measured at the start of the year-long baseline period for baseline observations and at the start of Program Year 1 for observations in the intervention period (Program Years 1 through 5). The estimates (and standard errors) in the impact estimate column show subgrou*p*-specific impacts, separately for each beneficiary characteristic listed in the table. We only tested differences *within* each subgroup if the estimates were significantly different *between* the two subgroups (that is, the *p*-value in the last column was <.10). Asterisks denote whether the impact estimate *within* a subgroup was significantly different from zero when estimates were significantly different between the subgroup categories. Because we could not observe diagnoses (which are used to determine HCCs and calculate HCC scores) at baseline for beneficiaries who were new to Medicare during the program years, we excluded new Medicare beneficiaries from all subgroup analyses (except the analysis based on dual status since beneficiaries who are new to Medicare, by definition, could not have been enrolled in both Medicare and Medicaid prior to joining Medicare). Due to this process, about 20 percent of observations from the regressions were excluded for the subgroups defined by HCC score and chronic conditions. Therefore, the main impact estimate of $1.1 PBPM for Track 1 overall may not lie between the impact estimates for these subgroups.

[a] We calculated percentage impacts relative to what the CPC+ mean would have been in Program Years 1 through 5 (separately and combined) in the absence of the intervention—that is, the unadjusted CPC+ mean minus the impact estimate.

[b] The *p*-values in the last column represent results from testing for statistically significant differences in impact estimates between the subgroups, based on the baseline beneficiary characteristic (using a t-test for all subgroups).

[c] The 12 frequently occurring chronic conditions are congestive heart failure, chronic obstructive pulmonary disease, history of acute myocardial infarction, ischemic heart disease, diabetes, metastatic cancer and acute leukemia, history of stroke, depression, dementia, atrial fibrillation, rheumatoid arthritis or osteoarthritis, and chronic kidney disease.

[d] For observations in the baseline year, hospitalizations are measured in 2015, the year before the start of the baseline year. For observations in the intervention period, hospitalizations are measured in 2016, the year before the start of Program Year 1.

*/**/*** Significantly different from zero at the 0.10/0.05/0.01 level, two-tailed test.

CMS = Centers for Medicare and Medicaid Services; HCC = hierarchical condition category; PBPM = per beneficiary per month; PY = Program Year.

Table 5.A.2.5b. Estimated average annual impacts of CPC+ on Medicare expenditures without CMS's enhanced payments across the five program years, by baseline beneficiary characteristics and SSP status, Track 1

Beneficiary subgroup definition, based on baseline characteristics	Track 1 – SSP				Track 1 – Non-SSP			
	Number (percentage) of CPC+ beneficiaries in subgroup at baseline	Impact estimate (standard error)	Percentage impact[a]	*p*-Value for difference in impact estimates between subgroups[a]	Number (percentage) of CPC+ beneficiaries in subgroup at baseline	Impact estimate (standard error)	Percentage impact[a]	*p*-Value for difference in impact estimates between subgroups
Main analysis (all beneficiaries)- average annual estimate for PY 1 through PY 5	-	**-$7.8* ($4.5)**	**-0.8%**	-	-	**$10.1** ($4.8)**	**1.1%**	-
Patients in the highest quartile of the HCC score distribution								
Yes	115,215 (26.8%)	-$22.4* ($12.3)	-1.1%		88,864 (25.0%)	$16.7 ($15.3)	0.8%	
No	315,425 (73.2%)	$0.2 ($4.0)	0.0%	0.07	266,666 (75.0%)	$8.5 ($3.9)	1.3%	0.60
Patients in the highest decile of the HCC score distribution or who have dementia								
Yes	68,759 (16.0%)	-$4.9 ($17.7)	-0.2%		54,382 (15.3%)	$5.9 ($20.5)	0.3%	
No	361,881 (84.0%)	-$6.3 ($4.1)	-0.8%	0.94	301,148 (84.7%)	$10.6 ($4.3)	1.4%	0.82
Patients with anxiety/depression or substance use disorders								
Yes	66,746 (16.8%)	-$5.1 ($12.7)	-0.4%		53,792 (16.4%)	$11.4 ($13.9)	0.9%	
No	329,703 (83.2%)	-$3.1 ($4.7)	-0.3%	0.88	273,568 (83.6%)	$10.6 ($4.9)	1.3%	0.96
Patients with multiple chronic conditions (at least 2 of 12 frequently occurring chronic conditions[c]) and one or more hospitalizations[d]								
Yes	38,153 (8.9%)	-$12.0 ($25.0)	-0.5%		30,089 (8.5%)	$19.6 ($28.5)	0.8%	
No	392,487 (91.1%)	-$5.3 ($4.4)	-0.6%	0.79	325,442 (91.5%)	$9.4 ($4.4)	1.1%	0.72
Patients dually eligible for Medicare and Medicaid								
Yes	55,728 (11.9%)	-$31.7** ($15.7)	-2.4%		51,626 (13.3%)	$4.9 ($14.7)	0.4%	
No	410,653 (88.1%)	-$4.7 ($4.4)	-0.5%	0.09	335,619 (86.7%)	$11.0 ($4.9)	1.2%	0.68

Source: Mathematica's analysis of Medicare claims data from January 2013 through December 2021.

Note: Beneficiary characteristics to determine subgroup membership are measured at the start of the year-long baseline period for baseline observations and at the start of Program Year 1 for observations in the intervention period (Program Years 1 through 5). The estimates (and standard errors) in the impact estimate column show subgroup*p*-specific impacts, separately for each beneficiary characteristic listed in the table. We only tested differences *within* each subgroup if the estimates were significantly different *between* the two subgroups (that is, the *p*-value in the last column was <.10). Asterisks denote whether the impact estimate *within* a subgroup was significantly different from zero when estimates were significantly different between the subgroup categories. Because we could not observe diagnoses (which are used to determine HCCs and calculate HCC scores) at baseline for beneficiaries who were new to Medicare during the program years, we excluded new Medicare beneficiaries from all subgroup analyses (except the analysis based on dual status since beneficiaries who are new to Medicare, by definition, could not have been enrolled in both Medicare and Medicaid prior to joining Medicare). Due to this process, about 20 percent of observations from the regressions were excluded for the subgroups defined by HCC score and chronic conditions. Therefore, the main impact estimate of -$7.8 PBPM for Track 1 SSP and $10.1 for Track 1 Non-SSP may not lie between the impact estimates for these subgroups.

[a] We calculated percentage impacts relative to what the CPC+ mean would have been in Program Years 1 through 5 (separately and combined) in the absence of the intervention—that is, the unadjusted CPC+ mean minus the impact estimate.

Table 5.A.2.5b. ***(continued)***

[b] The *p*-values in the last column represent results from testing for statistically significant differences in impact estimates between the subgroups, based on the baseline beneficiary characteristic (using a t-test for all subgroups).

[c] The 12 frequently occurring chronic conditions are congestive heart failure, chronic obstructive pulmonary disease, history of acute myocardial infarction, ischemic heart disease, diabetes, metastatic cancer and acute leukemia, history of stroke, depression, dementia, atrial fibrillation, rheumatoid arthritis or osteoarthritis, and chronic kidney disease.

[d] For observations in the baseline year, hospitalizations are measured in 2015, the year before the start of the baseline year. For observations in the intervention period, hospitalizations are measured in 2016, the year before the start of Program Year 1.

*/**/*** Significantly different from zero at the 0.10/0.05/0.01 level, two-tailed test.

CMS = Centers for Medicare and Medicaid Services; HCC = hierarchical condition category; PBPM = per beneficiary per month; PY = Program Year; SSP = Medicare Shared Savings Program.

Table 5.A.2.6a. Estimated average annual impacts of CPC+ on Medicare expenditures without CMS's enhanced payments across the five program years, by baseline beneficiary characteristics, Track 2

Beneficiary subgroup definition, based on baseline characteristics	Track 2 - Overall			
	Number (percentage) of CPC+ beneficiaries in subgroup at baseline	Impact estimate (standard error)	Percentage impact	*p*-Value for difference in impact estimates between subgroups[a]
Main analysis (all beneficiaries)- average annual estimate for PY 1 through PY 5	**-**	**$1.3 ($3.8)**	**0.1%**	**-**
Patients in the highest quartile of the HCC score distribution				
Yes	268,430 (26.1%)	$7.0 ($10.4)	0.3%	
No	761,970 (73.9%)	$5.0 ($3.1)	0.7%	0.85
Patients in the highest decile of the HCC score distribution or who have dementia				
Yes	162,510 (15.8%)	$9.4 ($14.6)	0.4%	
No	867,891 (84.2%)	$4.3 ($3.3)	0.5%	0.73
Patients with anxiety/depression or substance use disorders				
Yes	164,048 (17.3%)	$0.4 ($9.9)	0.0%	
No	784,877 (82.7%)	$4.5 ($3.6)	0.5%	0.67
Patients with multiple chronic conditions (at least 2 of 12 frequently occurring chronic conditions[b]) and one or more hospitalizations[c]				
Yes	90,543 (8.8%)	$20.5 ($19.4)	0.8%	
No	939,858 (91.2%)	$3.9 ($3.5)	0.4%	0.39
Patients dually eligible for Medicare and Medicaid				
Yes	140,782 (12.5%)	$6.1 ($11.0)	0.5%	
No	984,688 (87.5%)	$1.1 ($3.8)	0.1%	0.65

Source: Mathematica's analysis of Medicare claims data from January 2013 through December 2021.

Note: Beneficiary characteristics to determine subgroup membership are measured at the start of the year-long baseline period for baseline observations and at the start of Program Year 1 for observations in the intervention period (Program Years 1 through 5). The estimates (and standard errors) in the impact estimate column show subgrou*p*-specific impacts, separately for each beneficiary characteristic listed in the table. We only tested differences *within* each subgroup if the estimates were significantly different *between* the two subgroups (that is, the *p*-value in the last column was <.10). Asterisks denote whether the impact estimate *within* a subgroup was significantly different from zero when estimates were significantly different between the subgroup categories. Because we could not observe diagnoses (which are used to determine HCCs and calculate HCC scores) at baseline for beneficiaries who were new to Medicare during the program years, we excluded new Medicare beneficiaries from all subgroup analyses (except the analysis based on dual status since beneficiaries who are new to Medicare, by definition, could not have been enrolled in both Medicare and Medicaid prior to joining Medicare). Due to this process, about 20 percent of observations from the regressions were excluded for the subgroups defined by HCC score and chronic conditions. Therefore, the main impact estimate of $1.3 PBPM for Track 2 overall may not lie between the impact estimates for these subgroups.

[a] The *p*-values in the last column represent results from testing for statistically significant differences in impact estimates between the subgroups, based on the baseline beneficiary characteristic (using a t-test for all subgroups).

[b] The 12 frequently occurring chronic conditions are congestive heart failure, chronic obstructive pulmonary disease, history of acute myocardial infarction, ischemic heart disease, diabetes, metastatic cancer and acute leukemia, history of stroke, depression, dementia, atrial fibrillation, rheumatoid arthritis or osteoarthritis, and chronic kidney disease.

[c] For observations in the baseline year, hospitalizations are measured in 2015, the year before the start of the baseline year. For observations in the intervention period, hospitalizations are measured in 2016, the year before the start of Program Year 1.

*/**/*** Significantly different from zero at the 0.10/0.05/0.01 level, two-tailed test.

CMS = Centers for Medicare and Medicaid Services; HCC = hierarchical condition category; PBPM = per beneficiary per month; PY = Program Year.

Table 5.A.2.6b. Estimated average annual impacts of CPC+ on Medicare expenditures without CMS's enhanced payments across the five program years, by baseline beneficiary characteristics and SSP status, Track 2

Beneficiary subgroup definition, based on baseline characteristics	Track 2 – SSP				Track 2 – Non-SSP			
	Number (percentage) of CPC+ beneficiaries in subgroup at baseline	Impact estimate (standard error)	Percentage impact	*p*-Value for difference in impact estimates between subgroups[a]	Number (percentage) of CPC+ beneficiaries in subgroup at baseline	Impact estimate (standard error)	Percentage impact	*p*-Value for difference in impact estimates between subgroups[a]
Main analysis (all beneficiaries)- average annual estimate for PY 1 through PY 5	-	**-$8.1 ($6.1)**	**-0.8%**	-	-	**$8.9* ($4.7)**	**0.9%**	-
Patients in the highest quartile of the HCC score distribution								
Yes	120,947 (26.8%)	-$17.0 ($15.7)	-0.8%		146,522 (25.5%)	$26.0 ($13.5)	1.3%	
No	330,277 (73.2%)	$2.1 ($4.9)	0.3%	0.22	428,947 (74.5%)	$7.6 ($3.8)	1.1%	0.18
Patients in the highest decile of the HCC score distribution or who have dementia								
Yes	71,030 (15.7%)	-$28.9 ($22.4)	-1.2%		90,895 (15.8%)	$38.5** ($19.0)	1.7%	
No	380,194 (84.3%)	$1.4 ($5.3)	0.2%	0.18	484,574 (84.2%)	$6.9* ($4.1)	0.9%	0.10
Patients with anxiety/depression or substance use disorders								
Yes	74,382 (17.8%)	-$13.2 ($14.8)	-1.0%		89,058 (16.8%)	$10.9 ($13.1)	0.8%	
No	342,453 (82.2%)	-$2.1 ($5.6)	-0.2%	0.44	439,501 (83.2%)	$10.1 ($4.7)	1.2%	0.95
Patients with multiple chronic conditions (at least 2 of 12 frequently occurring chronic conditions[b]) and one or more hospitalizations[c]								
Yes	41,080 (9.1%)	$7.5 ($29.3)	0.3%		49,139 (8.5%)	$30.5 ($25.7)	1.2%	
No	410,144 (90.9%)	-$3.7 ($5.6)	-0.4%	0.70	526,331 (91.5%)	$10.1 ($4.3)	1.2%	0.43
Patients dually eligible for Medicare and Medicaid								
Yes	55,837 (11.3%)	-$12.8 ($17.6)	-0.9%		84,414 (13.5%)	$18.5 ($13.9)	1.5%	
No	438,154 (88.7%)	-$7.3 ($6.1)	-0.8%	0.75	542,895 (86.5%)	$8.0 ($4.7)	0.9%	0.45

Source: Mathematica's analysis of Medicare claims data from January 2013 through December 2021.

Note: Beneficiary characteristics to determine subgroup membership are measured at the start of the year-long baseline period for baseline observations and at the start of Program Year 1 for observations in the intervention period (Program Years 1 through 5). The estimates (and standard errors) in the impact estimate column show subgroup-specific impacts, separately for each beneficiary characteristic listed in the table. We only tested differences *within* each subgroup if the estimates were significantly different *between* the two subgroups (that is, the *p*-value in the last column was <.10). Asterisks denote whether the impact estimate *within* a subgroup was significantly different from zero when estimates were significantly different between the subgroup categories. Because we could not observe diagnoses (which are used to determine HCCs and calculate HCC scores) at baseline for beneficiaries who were new to Medicare during the program years, we excluded new Medicare beneficiaries from all subgroup analyses (except the analysis based on dual status since beneficiaries who are new to Medicare, by definition, could not have been enrolled in both Medicare and Medicaid prior to joining Medicare). Due to this process, about 20 percent of observations from the regressions were excluded for the subgroups defined by HCC score and chronic conditions. Therefore, the main impact estimate of -$8.1 PBPM for Track 2 SSP and $8.9 for Track 2 Non-SSP may not lie between the impact estimates for these subgroups.

[a] The *p*-values in the last column represent results from testing for statistically significant differences in impact estimates between the subgroups, based on the baseline beneficiary characteristic (using a t-test for all subgroups).

Table 5.A.2.6b. ***(continued)***

[b] The 12 frequently occurring chronic conditions are congestive heart failure, chronic obstructive pulmonary disease, history of acute myocardial infarction, ischemic heart disease, diabetes, metastatic cancer and acute leukemia, history of stroke, depression, dementia, atrial fibrillation, rheumatoid arthritis or osteoarthritis, and chronic kidney disease.

[c] For observations in the baseline year, hospitalizations are measured in 2015, the year before the start of the baseline year. For observations in the intervention period, hospitalizations are measured in 2016, the year before the start of Program Year 1.

*/**/*** Significantly different from zero at the 0.10/0.05/0.01 level, two-tailed test.

CMS = Centers for Medicare and Medicaid Services; HCC = hierarchical condition category; PBPM = per beneficiary per month; PY = Program Year; SSP = Medicare Shared Savings Program.

Table 5.A.2.7a. Estimated average annual impacts of CPC+ across the five program years on Medicare expenditures without CMS's enhanced payments, from main analysis and sensitivity tests, Track 1

Test	Motivation	Track 1 – Overall				
		Impact estimate	Percentage impact[a]	*p*-Value	90% CI lower bound	90% CI upper bound
Main analysis (average annual estimate for PY 1 through PY 5)	**Uses a difference-in-differences analysis with an ITT beneficiary sample, a one-year baseline period, controls for baseline beneficiary characteristics, COVID-19-related controls, and practice fixed effects**	**$1.1**	**0.1%**	**0.74**	**-$4.3**	**$6.6**
Altering length of baseline period						
Use two-year baseline period (instead of one year)[b]	Controls for outcome levels over longer pre-CPC+ period	$2.3	0.2%	0.44	-$2.6	$7.2
Altering the composition of the beneficiary sample						
Use sample of beneficiaries attributed during both the baseline and intervention periods as the analysis sample[c]	Helps to adjust for changes in sample composition between baseline and follow-up that may differ for the intervention and matched comparison groups	$1.3	0.1%	0.71	-$4.3	$6.9
Examine the impacts for the subset of beneficiaries attributed in the first quarter of the baseline period and the intervention period[d]	Removes any effects that may be due to changes in sample composition over time, for both baseline and intervention years	$6.6*	0.7%	0.05	$1.0	$12.1
Instead of following an ITT approach to defining the beneficiary sample (once attributed, beneficiaries stay in the sample for all subsequent years), allow beneficiaries to drop out of the sample if they no longer meet attribution requirements[e,f]	Assesses whether ITT tends to attenuate true effects by retaining beneficiaries in the intervention group who are no longer seen by CPC+ practices	$0.6	0.1%	0.86	-$4.8	$5.9
Altering the modeling assumptions						
Use generalized linear model with log link	Handles skewed expenditure distribution	-$3.0	-0.3%	0.57	-$11.5	$5.6
Trim expenditures at 98th percentile	Reduces influence of beneficiaries with high outlier expenditures	$0.2	0.0%	0.93	-$4.0	$4.5
Use log expenditures[g]	Reduces influence of beneficiaries with high outlier expenditures	NA	0.6%*	0.07	0.1%	1.2%
Use baseline beneficiary characteristics, practice characteristics, and practice-level averages of beneficiary characteristics (reflecting baseline characteristics of contemporaneous beneficiaries), all interacted with year indicators as additional controls (confounder test)	Accounts for potential time-varying effects of baseline beneficiary and practice characteristics on the outcome. Adjusts for practice-level measures of beneficiary characteristics to align with participation in CPC+ varying at the practice level	$7.3**	0.8%	0.02	$1.1	$13.6
Controlling for contemporaneous SSP participation						
Use a model that controls for contemporaneous (same year) SSP participation status	Controls for changes in SSP participation status among CPC+ and comparison practices over time	$1.7	0.2%	0.61	-$3.8	$7.1
Alternative definition of counterfactual						
Use a triple differences approach[h]	Controls for regional differences in trends among CPC+ and comparison practices	$6.5	0.7%	0.19	-$1.7	$14.6

Source: Mathematica's analysis of Medicare claims data from January 2013 through December 2021.

Table 5.A.2.7a. ***(continued)***

[a] We calculated percentage impacts relative to what the CPC+ mean would have been in Program Years 1 through 5 (separately and combined) in the absence of the intervention—that is, the unadjusted CPC+ mean minus the impact estimate.

[b] Sample size is 14 percent larger than the main analysis.

[c] Sample size is 35 percent smaller than the main analysis.

[d] Sample size is 32 percent smaller than the main analysis.

[e] Sample size is 11 percent smaller than the main analysis.

[f] The percentage of beneficiaries that are no longer attributed to CPC+ or comparison practices but are still included in the research sample due to the ITT approach grows over time; however, the yearly estimate from this sensitivity test was similar to the corresponding estimate from the main analysis in PY 5 (-$3.4 [$p$ = 0.49] and -$3.1 [$p$ = 0.51], respectively).

[g] We obtained only a percentage impact, not a dollar impact, from the model specification with log of expenditures as the outcome. The dollar magnitude of the impact in this model depends on the starting value—for example, a 0.8 percent impact for someone with expenditures equal to the CPC+ mean during the intervention period would be about $7.6.

[h] Sample size is 224 percent larger than the main analysis (because the triple-differences model also includes non-participating practices in CPC+ regions and unselected practices in comparison regions).

*/**/*** Significantly different from zero at the 0.10/0.05/0.01 level, two-tailed test.

NA = not applicable because the log expenditures test produces the difference-in-differences impact estimate in percentage term.

CI = confidence interval; CMS = Centers for Medicare and Medicaid Services; ITT = intent-to-treat; PY = Program Year.

Table 5.A.2.7b. Estimated average annual impacts of CPC+ across the five program years on Medicare expenditures without CMS's enhanced payments, from main analysis and sensitivity tests, Track 1 by SSP status

Test	Motivation	Track 1 – SSP					Track 1 – Non-SSP				
		Impact estimate	Percentage impact[a]	*p*-Value	90% CI lower bound	90% CI upper bound	Impact estimate	Percentage impact[a]	*p*-Value	90% CI lower bound	90% CI upper bound
Main analysis (average annual estimate for PY 1 through PY 5)	**Uses a difference-in-differences analysis with an ITT beneficiary sample, a one-year baseline period, controls for baseline beneficiary characteristics, COVID-19-related controls, and practice fixed effects**	**-$7.8***	**-0.8%**	**0.08**	**-$15.3**	**-$0.4**	**$10.1****	**1.1%**	**0.04**	**$2.1**	**$18.1**
Altering length of baseline period											
Use two-year baseline period (instead of one year)[b]	Controls for outcome levels over longer pre-CPC+ period	-$3.4	-0.3%	0.42	-$10.2	$3.5	$7.8*	0.8%	0.07	$0.7	$14.8
Altering the composition of the beneficiary sample											
Use sample of beneficiaries attributed during both the baseline and intervention periods as the analysis sample[c]	Helps to adjust for changes in sample composition between baseline and follow-up that may differ for the intervention and matched comparison groups	-$5.0	-0.5%	0.30	-$12.8	$2.9	$7.6	0.8%	0.13	-$0.6	$15.8
Examine the impacts for the subset of beneficiaries attributed in the first quarter of the baseline period and the intervention period[d]	Removes any effects that may be due to changes in sample composition over time, for both baseline and intervention years	$0.3	0.0%	0.95	-$7.3	$7.9	$12.8**	1.3%	0.01	$4.6	$21.0
Instead of following an ITT approach to defining the beneficiary sample (once attributed, beneficiaries stay in the sample for all subsequent years), allow beneficiaries to drop out of the sample if they no longer meet attribution requirements[e, f]	Assesses whether ITT tends to attenuate true effects by retaining beneficiaries in the intervention group who are no longer seen by CPC+ practices	-$7.5*	-0.8%	0.09	-$14.7	-$0.2	$9.1*	1.0%	0.06	$1.1	$17.1
Altering the modeling assumptions											
Use generalized linear model with log link	Handles skewed expenditure distribution	-$16.1**	-1.6%	0.02	-$27.5	-$4.7	$8.3	0.9%	0.27	-$4.2	$20.7
Trim expenditures at 98th percentile	Reduces influence of beneficiaries with high outlier expenditures	-$7.2**	-0.8%	0.04	-$13.1	-$1.4	$7.9**	0.9%	0.03	$1.9	$14.0
Use log expenditures[g]	Reduces influence of beneficiaries with high outlier expenditures	NA	-1.0%**	0.04	-1.9%	-0.2%	NA	2.4%***	0.00	1.6%	3.2%

Table 5.A.2.7b. ***(continued)***

Test	Motivation	Track 1 – SSP					Track 1 – Non-SSP				
		Impact estimate	Percentage impact[a]	*p*-Value	90% CI lower bound	90% CI upper bound	Impact estimate	Percentage impact[a]	*p*-Value	90% CI lower bound	90% CI upper bound
Use baseline beneficiary characteristics, practice characteristics, and practice-level averages of beneficiary characteristics (reflecting baseline characteristics of contemporaneous beneficiaries), all interacted with year indicators as additional controls (confounder test)	Accounts for potential time-varying effects of baseline beneficiary and practice characteristics on the outcome. Adjusts for practice-level measures of beneficiary characteristics to align with participation in CPC+ varying at the practice level	-$2.3	-0.2%	0.60	-$10.7	$6.1	$15.5***	1.7%	0.00	$6.2	$24.9
Controlling for contemporaneous SSP participation											
Use a model that controls for contemporaneous (same year) SSP participation status	Controls for changes in SSP participation status among CPC+ and comparison practices over time	-$3.4	-0.3%	0.46	-$10.8	$4.1	$11.8**	1.3%	0.01	$3.8	$19.7
Alternative definition of counterfactual											
Use a triple differences approach [h]	Controls for regional differences in trends among CPC+ and comparison practices	$3.4	0.3%	0.64	-$8.6	$15.5	$9.8	1.0%	0.14	-$1.0	$20.6

Source: Mathematica's analysis of Medicare claims data from January 2013 through December 2021.

[a] We calculated percentage impacts relative to what the CPC+ mean would have been in Program Years 1 through 5 (separately and combined) in the absence of the intervention—that is, the unadjusted CPC+ mean minus the impact estimate.

[b] Sample size is 14 percent larger than the main analysis.

[c] Sample size is about 35 percent smaller than the main analysis.

[d] Sample size is 32 percent smaller than the main analysis.

[e] Sample size is about 11 percent smaller than the main analysis.

[f] The percentage of beneficiaries that are no longer attributed to CPC+ or comparison practices but are still included in the research sample due to the ITT approach grows over time; however, the yearly estimate from this sensitivity test was similar to the corresponding estimate from the main analysis in PY 5 (-$16.1 [$p$ = 0.02] and -$19.5 [$p$ = 0.003] for Track 1 SSP and $8.1 [$p$ = 0.26] and $9.8 [$p$ = 0.15] for Track 1 Non-SSP, respectively).

[g] We obtained only a percentage impact, not a dollar impact, from the model specification with log of expenditures as the outcome. The dollar magnitude of the impact in this model depends on the starting value—for example, a 0.8 percent impact for someone with expenditures equal to the CPC+ mean during the intervention period would be about $7.6.

[h] Sample size is 129 to 348 percent larger than the main analysis (because the triple-differences model also includes non-participating practices in CPC+ regions and unselected practices in comparison regions).

*/**/*** Significantly different from zero at the 0.10/0.05/0.01 level, two-tailed test.

NA = not applicable because the log expenditures test produces the difference-in-differences impact estimate in percentage term.

CI = confidence interval; CMS = Centers for Medicare and Medicaid Services; ITT = intent-to-treat; PY = Program Year; SSP = Medicare Shared Savings Program.

Table 5.A.2.8a. Estimated average annual impacts of CPC+ across the five program years on Medicare expenditures without CMS's enhanced payments, from main analysis and sensitivity tests, Track 2

Test	Motivation	Track 2 – Overall				
		Impact estimate	Percentage impact	*p*-Value	90% CI lower bound	90% CI upper bound
Main analysis (average annual estimate for PY 1 through PY 5)	**Uses a difference-in-differences analysis with an ITT beneficiary sample, a one-year baseline period, controls for baseline beneficiary characteristics, COVID-19-related controls, and practice fixed effects**	**$1.3**	**0.1%**	**0.73**	**-$5.0**	**$7.7**
Altering length of baseline period						
Use two-year baseline period (instead of one year) [a]	Controls for outcome levels over longer pre-CPC+ period	$2.2	0.2%	0.52	-$3.3	$7.6
Altering the composition of the beneficiary sample						
Use sample of beneficiaries attributed during both the baseline and intervention periods as the analysis sample [b]	Helps to adjust for changes in sample composition between baseline and follow-up that may differ for the intervention and matched comparison groups	$3.0	0.3%	0.42	-$3.1	$9.0
Examine the impacts for the subset of beneficiaries attributed in the first quarter of the baseline period and the intervention period [c]	Removes any effects that may be due to changes in sample composition over time, for both baseline and intervention years	$4.3	0.4%	0.24	-$1.8	$10.4
Instead of following an ITT approach to defining the beneficiary sample (once attributed, beneficiaries stay in the sample for all subsequent years), allow beneficiaries to drop out of the sample if they no longer meet attribution requirements [d, e]	Assesses whether ITT tends to attenuate true effects by retaining beneficiaries in the intervention group who are no longer seen by CPC+ practices	$1.8	0.2%	0.64	-$4.6	$8.1
Altering the modeling assumptions						
Use generalized linear model with log link	Handles skewed expenditure distribution	-$0.8	-0.1%	0.89	-$10.5	$8.8
Trim expenditures at 98th percentile	Reduces influence of beneficiaries with high outlier expenditures	-$2.5	-0.3%	0.41	-$7.4	$2.5
Use log expenditures [f]	Reduces influence of beneficiaries with high outlier expenditures	NA	5.5%***	0.00	4.9%	6.2%
Use baseline beneficiary characteristics, practice characteristics, and practice-level averages of beneficiary characteristics (reflecting baseline characteristics of contemporaneous beneficiaries), all interacted with year indicators as additional controls (confounder test)	Accounts for potential time-varying effects of baseline beneficiary and practice characteristics on the outcome. Adjusts for practice-level measures of beneficiary characteristics to align with participation in CPC+ varying at the practice level	$9.5***	1.0%	0.01	$2.6	$16.4
Controlling for contemporaneous SSP participation						
Use a model that controls for contemporaneous (same year) SSP participation status	Controls for changes in SSP participation status among CPC+ and comparison practices over time	$1.6	0.2%	0.67	-$4.6	$7.9
Alternative definition of counterfactual						
Use a triple differences approach [g]	Controls for regional differences in trends among CPC+ and comparison practices	$5.7	0.6%	0.31	-$3.6	$14.9

Source: Mathematica's analysis of Medicare claims data from January 2013 through December 2021.

[a] Sample size is 14 percent larger than the main analysis.

Table 5.A.2.8a. ***(continued)***

[b] Sample size is 34 percent smaller than the main analysis.

[c] Sample size is 32 percent smaller than the main analysis.

[d] Sample size is 11 percent smaller than the main analysis.

[e] The percentage of beneficiaries that are no longer attributed to CPC+ or comparison practices but are still included in the research sample due to the ITT approach grows over time; however, the yearly estimate from this sensitivity test was similar to the corresponding estimate from the main analysis in PY 5 ($0.8 [$p$ = 0.89] and $0.7 [$p$ = 0.90], respectively).

[f] We obtained only a percentage impact, not a dollar impact, from the model specification with log of expenditures as the outcome. The dollar magnitude of the impact in this model depends on the starting value—for example, a 0.8 percent impact for someone with expenditures equal to the CPC+ mean during the intervention period would be about $7.6.

[g] Sample size is 225 percent larger than the main analysis (because the triple-differences model also includes non-participating practices in CPC+ regions and unselected practices in comparison regions).

*/**/*** Significantly different from zero at the 0.10/0.05/0.01 level, two-tailed test.

NA = not applicable because the log expenditures test produces the difference-in-differences impact estimate in percentage term.

CI = confidence interval; CMS = Centers for Medicare and Medicaid Services; ITT = intent-to-treat; PY = Program Year.

Table 5.A.2.8b. Estimated average annual impacts of CPC+ across the five program years on Medicare expenditures without CMS's enhanced payments, from main analysis and sensitivity tests, Track 2 by SSP status

Test	Motivation	Track 2 – SSP					Track 2 – Non-SSP				
		Impact estimate	Percentage impact	*p*-Value	90% CI lower bound	90% CI upper bound	Impact estimate	Percentage impact	*p*-Value	90% CI lower bound	90% CI upper bound
Main analysis (average annual estimate for PY 1 through PY 5)	**Uses a difference-in-differences analysis with an ITT beneficiary sample, a one-year baseline period, controls for baseline beneficiary characteristics, COVID-19-related controls, and practice fixed effects**	**-$8.1**	**-0.8%**	**0.18**	**-$18.1**	**$1.9**	**$8.9***	**0.9%**	**0.06**	**$1.2**	**$16.6**
Altering length of baseline period											
Use two-year baseline period (instead of one year) [a]	Controls for outcome levels over longer pre-CPC+ period	-$7.6	-0.8%	0.15	-$16.3	$1.0	$10.0**	1.1%	0.01	$3.4	$16.6
Altering the composition of the beneficiary sample											
Use sample of beneficiaries attributed during both the baseline and intervention periods as the analysis sample [b]	Helps to adjust for changes in sample composition between baseline and follow-up that may differ for the intervention and matched comparison groups	-$4.3	-0.4%	0.45	-$13.6	$5.0	$9.0*	0.9%	0.05	$1.4	$16.7
Examine the impacts for the subset of beneficiaries attributed in the first quarter of the baseline period and the intervention period [c]	Removes any effects that may be due to changes in sample composition over time, for both baseline and intervention years	-$3.1	-0.3%	0.58	-$12.3	$6.1	$10.5**	1.1%	0.03	$2.7	$18.3
Instead of following an ITT approach to defining the beneficiary sample (once attributed, beneficiaries stay in the sample for all subsequent years), allow beneficiaries to drop out of the sample if they no longer meet attribution requirements [d, e]	Assesses whether ITT tends to attenuate true effects by retaining beneficiaries in the intervention group who are no longer seen by CPC+ practices	-$5.4	-0.6%	0.38	-$15.4	$4.6	$7.8*	0.9%	0.10	$0.0	$15.5
Altering the modeling assumptions											
Use generalized linear model with log link	Handles skewed expenditure distribution	-$15.1*	-1.5%	0.09	-$29.5	-$0.7	$9.7	1.0%	0.20	-$2.8	$22.2
Trim expenditures at 98th percentile	Reduces influence of beneficiaries with high outlier expenditures	-$7.5	-0.8%	0.12	-$15.3	$0.4	$1.6	0.2%	0.66	-$4.4	$7.6
Use log expenditures [f]	Reduces influence of beneficiaries with high outlier expenditures	NA	3.8%***	0.00	2.8%	4.8%	NA	6.8%***	0.00	6.1%	7.6%

Table 5.A.2.8b. ***(continued)***

Test	Motivation	Track 2 – SSP					Track 2 – Non-SSP				
		Impact estimate	Percentage impact	*p*-Value	90% CI lower bound	90% CI upper bound	Impact estimate	Percentage impact	*p*-Value	90% CI lower bound	90% CI upper bound
Use baseline beneficiary characteristics, practice characteristics, and practice-level averages of beneficiary characteristics (reflecting baseline characteristics of contemporaneous beneficiaries), all interacted with year indicators as additional controls (confounder test)	Accounts for potential time-varying effects of baseline beneficiary and practice characteristics on the outcome. Adjusts for practice-level measures of beneficiary characteristics to align with participation in CPC+ varying at the practice level	-$0.9	-0.1%	0.87	-$11.1	$9.4	$17.0***	1.8%	0.00	$7.8	$26.3
Controlling for contemporaneous SSP participation											
Use a model that controls for contemporaneous (same year) SSP participation status	Controls for changes in SSP participation status among CPC+ and comparison practices over time	-$3.4	-0.3%	0.56	-$13.3	$6.4	$10.1**	1.1%	0.03	$2.3	$17.9
Alternative definition of counterfactual											
Use a triple differences approach [g]	Controls for regional differences in trends among CPC+ and comparison practices	-$0.3	0.0%	0.97	-$15.8	$15.2	$9.1	1.0%	0.15	-$1.2	$19.4

Source: Mathematica's analysis of Medicare claims data from January 2013 through December 2021.

[a] Sample size is 14 percent larger than the main analysis.

[b] Sample size is 34 percent smaller than the main analysis.

[c] Sample size is 32 percent smaller than the main analysis.

[d] Sample size is about 11 percent smaller than the main analysis.

[e] The percentage of beneficiaries that are no longer attributed to CPC+ or comparison practices but are still included in the research sample due to the ITT approach grows over time; however, the yearly estimate from this sensitivity test was similar to the corresponding estimate from the main analysis in PY 5 (-$11.5 [$p$ = 0.24] and -$17.4 [$p$ = 0.06] for Track 2 SSP and $9.8 [$p$ = 0.17] and $14.1 [$p$ = 0.03] for Track 2 Non-SSP, respectively).

[f] We obtained only a percentage impact, not a dollar impact, from the model specification with log of expenditures as the outcome. The dollar magnitude of the impact in this model depends on the starting value—for example, a 0.8 percent impact for someone with expenditures equal to the CPC+ mean during the intervention period would be about $7.6.

[g] Sample size is 155 to 290 percent larger than the main analysis (because the triple-differences model also includes non-participating practices in CPC+ regions and unselected practices in comparison regions).

*/**/*** Significantly different from zero at the 0.10/0.05/0.01 level, two-tailed test.

NA = not applicable because the log expenditures test produces the difference-in-differences impact estimate in percentage term.

CI = confidence interval; CMS = Centers for Medicare and Medicaid Services; ITT = intent-to-treat; PY = Program Year; SSP = Medicare Shared Savings Program.

Table 5.A.2.9a. Estimated impacts of CPC+ on Medicare expenditures without CMS's enhanced payments in PYs 4 and 5, from main analysis and COVID-19 specific sensitivity tests, Track 1

	Track 1 – Overall				
Year	Impact estimate	Percentage impact[a]	*p*-Value	90% CI lower bound	90% CI upper bound
Main analysis that uses a difference-in-differences empirical strategy					
PY 4 estimate	-$2.8	-0.3%	0.54	-$10.2	$4.7
PY 5 estimate	-$3.1	-0.3%	0.51	-$10.9	$4.6
Triple Differences Approach that controls for regional differences in trends due to COVID-19 among CPC+ and comparison practices[b]					
PY 4 estimate	$4.2	0.4%	0.54	-$7.0	$15.5
PY 5 estimate	$5.3	0.5%	0.48	-$7.1	$17.6
Estimates obtained for outcome constructed by dropping claims from March 2020 to May 2020 (to test for sensitivity to change in expenditures at the start of the pandemic)[c]					
PY 4 estimate	-$0.5	0.0%	0.92	-$8.0	$7.1

Source: Mathematica's analysis of Medicare claims data from January 2013 through December 2021.

[a] We calculated percentage impacts relative to what the CPC+ mean would have been in Program Years 1 through 5 (separately and combined) in the absence of the intervention—that is, the unadjusted CPC+ mean minus the impact estimate.

[b] Sample size is 224 percent larger than the main analysis (because the triple-differences model also includes non-participating practices in CPC+ regions and unselected practices in comparison regions).

[c] Sample size is about 0.01 percent smaller than the main analysis.

*/**/*** Significantly different from zero at the 0.10/0.05/0.01 level, two-tailed test.

CI = confidence interval; CMS = Centers for Medicare and Medicaid Services; PY = Program Year.

Table 5.A.2.9b. Estimated impacts of CPC+ on Medicare expenditures without CMS's enhanced payments in PYs 4 and 5, from main analysis and COVID-19 specific sensitivity tests, Track 1 by SSP status

	Track 1 – SSP					Track 1 – Non-SSP				
Year	Impact estimate	Percentage impact[a]	*p*-Value	90% CI lower bound	90% CI upper bound	Impact estimate	Percentage impact[a]	*p*-Value	90% CI lower bound	90% CI upper bound
Main analysis that uses a difference-in-differences empirical strategy										
PY 4 estimate	-$15.1**	-1.5%	0.02	-$25.9	-$4.4	$10.0	1.1%	0.11	-$0.3	$20.2
PY 5 estimate	-$19.5***	-1.8%	0.00	-$30.3	-$8.6	$9.8	1.0%	0.15	-$1.3	$20.9
Triple Differences Approach that controls for regional differences in trends due to COVID-19 among CPC+ and comparison practices[b]										
PY 4 estimate	-$2.6	-0.3%	0.80	-$19.3	$14.1	$12.5	1.4%	0.16	-$2.2	$27.2
PY 5 estimate	-$10.7	-1.0%	0.34	-$29.2	$7.8	$20.7**	2.1%	0.03	$4.6	$36.8
Estimates obtained for outcome constructed by dropping claims from March 2020 to May 2020 (to test for sensitivity to change in expenditures at the start of the pandemic)[c]										
PY 4 estimate	-$11.9*	-1.1%	0.07	-$22.7	-$1.1	$10.8*	1.1%	0.10	$0.1	$21.5

Source: Mathematica's analysis of Medicare claims data from January 2013 through December 2021.

[a] We calculated percentage impacts relative to what the CPC+ mean would have been in Program Years 1 through 5 (separately and combined) in the absence of the intervention—that is, the unadjusted CPC+ mean minus the impact estimate.

[b] Sample size is 129 to 348 percent larger than the main analysis (because the triple-differences model also includes non-participating practices in CPC+ regions and unselected practices in comparison regions).

[c] Sample size is about 0.01 percent smaller than the main analysis.

*/**/*** Significantly different from zero at the 0.10/0.05/0.01 level, two-tailed test.

CI = confidence interval; CMS = Centers for Medicare and Medicaid Services; PY = Program Year.

Table 5.A.2.10a. Estimated impacts of CPC+ on Medicare expenditures without CMS's enhanced payments in PYs 4 and 5, from main analysis and COVID-19 specific sensitivity tests, Track 2

	Track 2 – Overall				
Year	Impact estimate	Percentage impact	*p*-Value	90% CI lower bound	90% CI upper bound
Main analysis that uses a difference-in-differences empirical strategy					
PY 4 estimate	-$2.3	-0.2%	0.65	-$10.8	$6.1
PY 5 estimate	$0.7	0.1%	0.90	-$8.5	$9.9
Triple Differences Approach that controls for regional differences in trends due to COVID-19 among CPC+ and comparison practices[a]					
PY 4 estimate	$3.4	0.4%	0.65	-$9.1	$16.0
PY 5 estimate	$10.6	1.0%	0.20	-$3.0	$24.1
Estimates obtained for outcome constructed by dropping claims from March 2020 to May 2020 (to test for sensitivity to change in expenditures at the start of the pandemic)[b]					
PY 4 estimate	-$1.8	-0.2%	0.73	-$10.5	$6.9

Source: Mathematica's analysis of Medicare claims data from January 2013 through December 2021.

[a] Sample size is 225 percent larger than the main analysis (because the triple-differences model also includes non-participating practices in CPC+ regions and unselected practices in comparison regions).

[b] Sample size is 0.01 percent smaller than the main analysis.

*/**/*** Significantly different from zero at the 0.10/0.05/0.01 level, two-tailed test.

CI = confidence interval; CMS = Centers for Medicare and Medicaid Services; PY = Program Year.

Table 5.A.2.10b. Estimated impacts of CPC+ on Medicare expenditures without CMS's enhanced payments in PYs 4 and 5, from main analysis and COVID-19 specific sensitivity tests, Track 2 by SSP status

	Track 2– SSP					Track 2 – Non-SSP				
Year	Impact estimate	Percentage impact	*p*-Value	90% CI lower bound	90% CI upper bound	Impact estimate]	Percentage impact	*p*-Value	90% CI lower bound	90% CI upper bound
Main analysis that uses a difference-in-differences empirical strategy										
PY 4 estimate	-$14.3*	-1.5%	0.08	-$27.8	-$0.9	$9.0	1.0%	0.14	-$0.9	$18.8
PY 5 estimate	-$17.4*	-1.6%	0.06	-$32.4	-$2.5	$14.1**	1.4%	0.03	$3.2	$25.1
Triple Differences Approach that controls for regional differences in trends due to COVID-19 among CPC+ and comparison practices[a]										
PY 4 estimate	-$2.9	-0.3%	0.82	-$23.5	$17.7	$9.7	1.0%	0.25	-$4.1	$23.6
PY 5 estimate	-$8.4	-0.8%	0.55	-$31.8	$15.0	$21.8**	2.2%	0.02	$7.0	$36.5
Estimates obtained for outcome constructed by dropping claims from March 2020 to May 2020 (to test for sensitivity to change in expenditures at the start of the pandemic)[b]										
PY 4 estimate	-$14.5*	-1.4%	0.08	-$28.4	-$0.7	$9.5	1.0%	0.13	-$0.7	$19.6

Source: Mathematica's analysis of Medicare claims data from January 2013 through December 2021.

[a] Sample size is 155 to 290 percent larger than the main analysis (because the triple-differences model also includes non-participating practices in CPC+ regions and unselected practices in comparison regions).

[b] Sample size is 0.01 percent smaller than the main analysis.

*/**/*** Significantly different from zero at the 0.10/0.05/0.01 level, two-tailed test.

CI = confidence interval; CMS = Centers for Medicare and Medicaid Services; PY = Program Year; SSP = Medicare Shared Savings Program.

5.A.3. Claims-based quality of care measures

Table 5.A.3.1a. Regression-adjusted means and estimated impacts of CPC+ on selected claims-based quality-of-care measures for attributed Medicare FFS beneficiaries by program year and average across the five program years, Track 1

	Track 1—Overall				
	CPC+ mean[a]	C mean[a]	Impact estimate[b] (SE)	90% confidence interval	*p*-Value
Planned care and population health measures for beneficiaries ages 18–75 with diabetes (percentage)					
Received HbA1c test					
Baseline	90.8%	91.6%	NA	NA	NA
PY 1	91.1%	91.9%	0.0 (0.2)	(-0.2, 0.3)	0.75
PY 2	91.2%	91.8%	0.1 (0.2)	(-0.2, 0.4)	0.51
PY 3	91.3%	91.7%	0.4* (0.2)	(0.0, 0.8)	0.08
PY 4	88.5%	88.6%	0.6** (0.3)	(0.2, 1.1)	0.01
PY 5	90.6%	91.1%	0.3 (0.3)	(-0.2, 0.8)	0.34
PY 1 through 5	90.5%	91.0%	0.3* (0.2)	(0.0, 0.6)	0.10
Received eye exam					
Baseline	63.5%	64.4%	NA	NA	NA
PY 1	64.8%	65.0%	0.7*** (0.2)	(0.3, 1.1)	0.00
PY 2	65.7%	65.3%	1.3*** (0.3)	(0.9, 1.7)	0.00
PY 3	65.6%	65.9%	0.6** (0.3)	(0.1, 1.1)	0.04
PY 4	61.4%	61.2%	1.1*** (0.3)	(0.6, 1.7)	0.00
PY 5	63.6%	63.6%	0.9*** (0.3)	(0.3, 1.5)	0.01
PY 1 through 5	64.2%	64.2%	0.9*** (0.2)	(0.5, 1.3)	0.00
Received attention for nephropathy					
Baseline	80.9%	80.9%	NA	NA	NA
PY 1	81.9%	81.2%	0.7*** (0.2)	(0.3, 1.0)	0.01
PY 2	82.4%	81.3%	1.1*** (0.3)	(0.6, 1.5)	0.00
PY 3	82.4%	81.6%	0.7** (0.3)	(0.2, 1.3)	0.03
PY 4	79.0%	78.3%	0.7* (0.4)	(0.1, 1.3)	0.07
PY 5	81.1%	81.0%	0.1 (0.4)	(-0.6, 0.7)	0.83
PY 1 through 5	81.3%	80.7%	0.7** (0.3)	(0.2, 1.1)	0.02
Diabetes Composite Measure 1 (received all three tests above: HbA1c test, eye exam, attention for nephropathy)					
Baseline	51.0%	51.9%	NA	NA	NA
PY 1	52.4%	52.8%	0.6** (0.3)	(0.1, 1.0)	0.04
PY 2	53.7%	53.0%	1.6*** (0.3)	(1.1, 2.1)	0.00
PY 3	53.6%	53.6%	0.9** (0.4)	(0.3, 1.5)	0.02
PY 4	48.0%	47.4%	1.6*** (0.4)	(0.9, 2.2)	0.00

Table 5.A.3.1a. ***(continued)***

	Track 1—Overall				
	CPC+ mean[a]	C mean[a]	Impact estimate[b] (SE)	90% confidence interval	*p*-Value
PY 5	51.1%	51.1%	0.9** (0.4)	(0.2, 1.6)	0.03
PY 1 through 5	51.8%	51.5%	1.1*** (0.3)	(0.6, 1.6)	0.00
Diabetes Composite Measure 2 (received none of the three tests above)					
Baseline	2.5%	2.3%	NA	NA	NA
PY 1	2.3%	2.3%	-0.2** (0.1)	(-0.3, -0.1)	0.01
PY 2	2.3%	2.3%	-0.2** (0.1)	(-0.4, -0.1)	0.01
PY 3	2.3%	2.2%	-0.2* (0.1)	(-0.3, 0.0)	0.05
PY 4	3.5%	3.5%	-0.2* (0.1)	(-0.4, 0.0)	0.07
PY 5	2.5%	2.4%	-0.1 (0.1)	(-0.2, 0.1)	0.33
PY 1 through 5	2.6%	2.5%	-0.2** (0.1)	(-0.3, -0.1)	0.02
Unweighted sample sizes for the diabetes measures[c]					
Number of beneficiaries	266,315	922,508			
Number of beneficiary-years	789,897	2,700,565			
Planned care and population health measures for female beneficiaries ages 52–74 (percentage)					
Received breast cancer screening					
Baseline	72.6%	73.2%	NA	NA	NA
PY 1	73.5%	73.7%	0.4*** (0.2)	(0.2, 0.7)	0.01
PY 2	74.3%	74.0%	0.9*** (0.2)	(0.6, 1.3)	0.00
PY 3	74.9%	74.7%	0.8*** (0.2)	(0.4, 1.2)	0.00
PY 4	73.0%	72.9%	0.7*** (0.2)	(0.3, 1.1)	0.01
PY 5	73.3%	73.0%	0.8*** (0.3)	(0.4, 1.3)	0.00
PY 1 through 5	73.8%	73.7%	0.7*** (0.2)	(0.4, 1.1)	0.00
Unweighted sample sizes for the breast cancer screening measure[c]					
Number of beneficiaries	440,433	1,487,754			
Number of beneficiary-years	1,357,359	4,546,561			
Planned care and population health measures for beneficiaries ages 21 and older[d]					
Percentage of beneficiaries with cardiovascular disease who were prescribed and filled statin therapy					
Baseline	58.9%	59.1%	NA	NA	NA
PY 1	60.2%	60.4%	0.0 (0.1)	(-0.2, 0.1)	0.78
PY 2	59.4%	59.8%	-0.2 (0.1)	(-0.4, 0.0)	0.17
PY 3	60.7%	61.0%	-0.2 (0.2)	(-0.4, 0.1)	0.27
PY 4	61.3%	61.8%	-0.3* (0.2)	(-0.6, 0.0)	0.09
PY 5	63.0%	63.4%	-0.3 (0.2)	(-0.6, 0.1)	0.18
PY 1 through 5	61.0%	61.3%	-0.2 (0.1)	(-0.4, 0.0)	0.16

Table 5.A.3.1a. ***(continued)***

	Track 1—Overall				
	CPC+ mean[a]	C mean[a]	Impact estimate[b] (SE)	90% confidence interval	*p*-Value
Unweighted sample sizes for the statin therapy measure[c]					
Number of beneficiaries	837,030	2,810,058			
Number of beneficiary-years	2,816,058	9,347,679			
Planned care and population health measures for beneficiaries ages 18 and older[d]					
Percentage of beneficiaries on diabetes medications with proportion of days covered by medication > 80%					
Baseline	77.2%	77.2%	NA	NA	NA
PY 1	78.2%	78.6%	-0.3 (0.2)	(-0.7, 0.1)	0.16
PY 2	79.2%	79.5%	-0.3 (0.2)	(-0.7, 0.1)	0.29
PY 3	78.7%	79.1%	-0.3 (0.3)	(-0.8, 0.1)	0.17
PY 4	82.1%	82.0%	0.2 (0.3)	(-0.3, 0.6)	0.51
PY 5	80.9%	81.4%	-0.5* (0.3)	(-0.9, 0.0)	0.09
PY 1 through 5	79.9%	80.2%	-0.2 (0.2)	(-0.6, 0.1)	0.22
Percentage of beneficiaries on renin-angiotensin system antagonists with proportion of days covered by medication > 80%					
Baseline	78.8%	78.6%	NA	NA	NA
PY 1	81.0%	81.0%	-0.3* (0.1)	(-0.5, 0.0)	0.06
PY 2	82.0%	81.8%	-0.1 (0.1)	(-0.3, 0.1)	0.39
PY 3	81.9%	82.2%	-0.5*** (0.1)	(-0.8, -0.3)	0.00
PY 4	84.0%	84.2%	-0.4*** (0.1)	(-0.7, -0.2)	0.00
PY 5	83.7%	84.1%	-0.7*** (0.1)	(-0.9, -0.4)	0.00
PY 1 through 5	82.6%	82.8%	-0.4*** (0.1)	(-0.6, -0.2)	0.00
Percentage of beneficiaries on statins with proportion of days covered by medication > 80%					
Baseline	76.0%	76.3%	NA	NA	NA
PY 1	76.0%	76.3%	-0.1 (0.1)	(-0.3, 0.2)	0.73
PY 2	79.4%	79.7%	0.0 (0.1)	(-0.2, 0.3)	0.72
PY 3	79.5%	79.8%	0.0 (0.1)	(-0.3, 0.2)	0.77
PY 4	82.5%	82.8%	-0.1 (0.2)	(-0.3, 0.2)	0.62
PY 5	81.9%	82.3%	-0.2 (0.2)	(-0.4, 0.1)	0.23
PY 1 through 5	80.0%	80.4%	-0.1 (0.1)	(-0.3, 0.1)	0.62
Percentage of beneficiaries with both coronary artery disease (CAD) and diabetes who were prescribed and filled angiotensin-converting enzyme (ACE) inhibitor or angiotensin receptor blocker (ARB) therapy					
Baseline	78.2%	77.7%	NA	NA	NA
PY 1	78.3%	77.6%	0.2 (0.3)	(-0.2, 0.6)	0.46
PY 2	77.7%	77.2%	0.1 (0.3)	(-0.4, 0.6)	0.75
PY 3	77.8%	77.3%	0.1 (0.3)	(-0.4, 0.6)	0.78

Table 5.A.3.1a. ***(continued)***

	Track 1—Overall				
	CPC+ mean[a]	C mean[a]	Impact estimate[b] (SE)	90% confidence interval	*p*-Value
PY 4	77.0%	76.4%	0.2 (0.3)	(-0.4, 0.7)	0.66
PY 5	77.3%	76.6%	0.2 (0.3)	(-0.4, 0.7)	0.59
PY 1 through 5	77.6%	77.0%	0.1 (0.3)	(-0.3, 0.6)	0.59
Unweighted sample sizes for percentage of beneficiaries on diabetes medications with proportion of days covered by medication > 80%					
Number of beneficiaries	188,857	641,396			
Number of beneficiary-years	566,318	1,908,338			
Unweighted sample sizes for percentage of beneficiaries on renin-angiotensin system antagonists with proportion of days covered by medication > 80%					
Number of beneficiaries	565,577	1,916,159			
Number of beneficiary-years	1,775,795	5,956,372			
Unweighted sample sizes for percentage of beneficiaries on statins with proportion of days covered by medication > 80%					
Number of beneficiaries	651,937	2,217,912			
Number of beneficiary-years	2,128,767	7,197,407			
Unweighted sample sizes for percentage of beneficiaries with both CAD and diabetes who were prescribed and filled ACE inhibitor or ARB therapy					
Number of beneficiaries	158,692	503,445			
Number of beneficiary-years	351,130	1,103,282			
Measures for continuity of care[e]					
Percentage of primary care ambulatory visits at assigned practice					
Baseline	75.5%	73.7%	NA	NA	NA
PY 1	72.5%	70.6%	0.1 (0.2)	(-0.4, 0.5)	0.83
PY 2	64.0%	61.7%	0.4 (0.5)	(-0.4, 1.2)	0.40
PY 3	61.4%	58.7%	0.8 (0.5)	(0.0, 1.7)	0.10
PY 4	54.5%	51.7%	1.0 (0.8)	(-0.3, 2.2)	0.21
PY 5	50.4%	48.1%	0.4 (0.7)	(-0.8, 1.7)	0.54
PY 1 through 5	60.1%	57.7%	0.5 (0.5)	(-0.2, 1.3)	0.24
Across all PCPs and specialists providing care to a patient, where each practitioner in the beneficiary's assigned practice is treated separately					
Percentage of visits with the usual provider of care (UPC)					
Baseline	48.4%	48.4%	NA	NA	NA
PY 1	47.4%	47.4%	-0.1 (0.1)	(-0.2, 0.0)	0.30
PY 2	46.2%	46.3%	-0.1* (0.1)	(-0.3, 0.0)	0.09
PY 3	45.5%	45.5%	-0.1 (0.1)	(-0.2, 0.1)	0.36
PY 4	47.7%	48.1%	-0.4*** (0.1)	(-0.6, -0.2)	0.00
PY 5	44.9%	45.2%	-0.2** (0.1)	(-0.4, -0.1)	0.02
PY 1 through 5	46.3%	46.5%	-0.2** (0.1)	(-0.3, -0.1)	0.02

Table 5.A.3.1a. ***(continued)***

	Track 1—Overall				
	CPC+ mean[a]	C mean[a]	Impact estimate[b] (SE)	90% confidence interval	*p*-Value
Reversed Bice-Boxerman fragmentation of care index					
Baseline	76.9%	77.2%	NA	NA	NA
PY 1	77.9%	78.2%	0.0 (0.1)	(-0.1, 0.1)	0.83
PY 2	79.1%	79.3%	0.1 (0.1)	(-0.1, 0.2)	0.44
PY 3	79.8%	80.1%	0.0 (0.1)	(-0.2, 0.1)	0.77
PY 4	80.1%	80.3%	0.1 (0.1)	(-0.1, 0.2)	0.55
PY 5	81.3%	81.4%	0.2 (0.1)	(0.0, 0.3)	0.17
PY 1 through 5	79.7%	79.9%	0.1 (0.1)	(-0.1, 0.2)	0.51
Across all PCPs and specialists providing care to a patient, where all practitioners in the beneficiary's assigned practice are treated as a single practitioner					
Percentage of visits with the usual provider of care (UPC)					
Baseline	51.0%	51.0%	NA	NA	NA
PY 1	49.9%	50.0%	-0.1 (0.1)	(-0.2, 0.0)	0.17
PY 2	48.1%	48.3%	-0.2** (0.1)	(-0.3, 0.0)	0.04
PY 3	48.0%	48.1%	-0.1 (0.1)	(-0.2, 0.1)	0.54
PY 4	49.7%	50.0%	-0.3* (0.1)	(-0.5, 0.0)	0.05
PY 5	46.9%	47.1%	-0.2 (0.1)	(-0.4, 0.1)	0.22
PY 1 through 5	48.5%	48.7%	-0.2* (0.1)	(-0.3, 0.0)	0.08
Reversed Bice-Boxerman fragmentation of care index					
Baseline	74.1%	74.3%	NA	NA	NA
PY 1	75.2%	75.4%	0.0 (0.1)	(-0.1, 0.2)	0.60
PY 2	77.0%	77.1%	0.1 (0.1)	(0.0, 0.3)	0.16
PY 3	77.1%	77.4%	0.0 (0.1)	(-0.2, 0.2)	0.77
PY 4	78.0%	78.2%	0.0 (0.2)	(-0.3, 0.2)	0.91
PY 5	79.3%	79.5%	0.1 (0.1)	(-0.2, 0.3)	0.59
PY 1 through 5	77.4%	77.6%	0.0 (0.1)	(-0.1, 0.2)	0.67
Unweighted sample sizes for percentage of primary care ambulatory visits at assigned practice[c]					
Number of beneficiaries	1,329,436	4,561,521			
Number of beneficiary-years	4,672,381	15,847,830			
Unweighted sample sizes for percentage of visits with the usual provider of care[c]					
Number of beneficiaries	1,359,497	4,669,754			
Number of beneficiary-years	4,932,540	16,728,879			
Unweighted sample sizes for reversed Bice-Boxerman fragmentation of care index[c]					
Number of beneficiaries	1,239,589	4,235,238			
Number of beneficiary-years	4,056,189	13,656,331			

Table 5.A.3.1a. ***(continued)***

	Track 1—Overall				
	CPC+ mean[a]	C mean[a]	Impact estimate[b] (SE)	90% confidence interval	*p*-Value
Comprehensiveness of care (measured at the physician level)					
Physician involvement in patient conditions[f]					
Baseline	67.8%	68.1%	NA	NA	NA
PY 1	68.0%	67.9%	0.4* (0.3)	(0.0, 0.9)	0.09
PY 2	68.3%	67.9%	0.6** (0.3)	(0.1, 1.1)	0.04
PY 3	69.2%	68.8%	0.7** (0.3)	(0.1, 1.3)	0.04
PY 4	71.6%	71.1%	0.8** (0.4)	(0.2, 1.4)	0.04
PY 5	69.1%	68.8%	0.6 (0.4)	(0.0, 1.3)	0.12
PY 1 through 5	69.1%	68.8%	0.6** (0.3)	(0.2, 1.0)	0.02
Range of services provided by physicians[h]					
Baseline	2.08	1.93	NA	NA	NA
PY 1	2.03	1.89	0.00 (0.02)	(-0.03, 0.03)	0.98
PY 2	1.99	1.84	0.01 (0.02)	(-0.02, 0.04)	0.55
PY 3	1.95	1.84	-0.03 (0.02)	(-0.06, 0.01)	0.24
PY 4	1.86	1.78	-0.05** (0.02)	(-0.09, -0.01)	0.03
PY 5	1.80	1.70	-0.04 (0.03)	(-0.09, 0.00)	0.10
PY 1 through 5	1.94	1.81	-0.02 (0.02)	(-0.04, 0.01)	0.31
Management of new problems by physicians[g]					
Baseline	1.001	1.002	NA	NA	NA
PY 1	1.001	1.001	0.001 (0.002)	(-0.002, 0.004)	0.66
PY 2	1.000	1.000	0.001 (0.002)	(-0.002, 0.005)	0.63
PY 3	1.000	1.001	0.000 (0.002)	(-0.004, 0.004)	0.95
PY 4	1.000	1.000	0.001 (0.003)	(-0.004, 0.005)	0.84
PY 1 through 4	1.000	1.000	0.001 (0.002)	(-0.002, 0.003)	0.72
Unweighted sample sizes for physician involvement in patient conditions					
Number of physicians	4,176	15,854			
Number of physician-years	20,029	73,999			
Unweighted sample sizes for range of services provided by physicians					
Number of physicians	4,063	15,151			
Number of physician-years	19,446	69,700			
Unweighted sample sizes for management of new problems by physicians					
Number of physicians	4,102	15,511			
Number of physician-years	17,049	62,654			

Table 5.A.3.1a. ***(continued)***

	Track 1—Overall				
	CPC+ mean[a]	C mean[a]	Impact estimate[b] (SE)	90% confidence interval	*p*-Value
Patient and caregiver engagement					
Percentage of beneficiaries receiving hospice services					
Baseline	2.7%	2.7%	NA	NA	NA
PY 1	2.8%	2.7%	0.1* (0.0)	(0.0, 0.1)	0.08
PY 2	2.9%	2.8%	0.1** (0.0)	(0.0, 0.1)	0.02
PY 3	3.1%	2.9%	0.1*** (0.0)	(0.1, 0.2)	0.00
PY 4	3.3%	3.2%	0.1** (0.0)	(0.0, 0.1)	0.02
PY 5	3.3%	3.1%	0.1*** (0.0)	(0.1, 0.2)	0.00
PY 1 through 5	3.1%	3.0%	0.1*** (0.0)	(0.1, 0.1)	0.00
Length of hospice stay, in days (for beneficiaries receiving hospice services)					
Baseline	60	65	NA	NA	NA
PY 1	62	66	1.6 (1.0)	(-0.1, 3.3)	0.11
PY 2	66	69	2.7** (1.1)	(0.9, 4.5)	0.01
PY 3	71	73	3.2*** (1.1)	(1.4, 5.0)	0.00
PY 4	69	70	3.6*** (1.2)	(1.6, 5.6)	0.00
PY 5	70	71	4.2*** (1.2)	(2.3, 6.2)	0.00
PY 1 through 5	68	70	3.1*** (0.9)	(1.6, 4.6)	0.00
Length of hospice stay, in days (for all beneficiaries)					
Baseline	1.6	1.8	NA	NA	NA
PY 1	1.7	1.8	0.1*** (0.0)	(0.0, 0.1)	0.01
PY 2	1.9	1.9	0.1*** (0.0)	(0.1, 0.2)	0.00
PY 3	2.2	2.1	0.2*** (0.0)	(0.1, 0.3)	0.00
PY 4	2.3	2.2	0.2*** (0.0)	(0.1, 0.3)	0.00
PY 5	2.3	2.2	0.3*** (0.0)	(0.2, 0.3)	0.00
PY 1 through 5	2.1	2.1	0.2*** (0.0)	(0.1, 0.2)	0.00
Unweighted sample sizes for patient and caregiver engagement measures					
Number of beneficiaries for length of hospice stay	141,943	455,931			
Other quality measures					
Percentage of index acute care hospital discharges that were followed by an unplanned readmission within 30 days					
Baseline	15.5%	15.7%	NA	NA	NA
PY 1	15.7%	15.8%	0.1 (0.1)	(-0.1, 0.4)	0.38
PY 2	15.8%	15.9%	0.2 (0.2)	(0.0, 0.5)	0.12
PY 3	15.8%	16.1%	0.0 (0.2)	(-0.2, 0.3)	0.87
PY 4	16.3%	16.1%	0.4*** (0.2)	(0.2, 0.7)	0.01
PY 5	16.1%	16.3%	0.1 (0.2)	(-0.2, 0.4)	0.61

Table 5.A.3.1a. ***(continued)***

	Track 1—Overall				
	CPC+ mean[a]	C mean[a]	Impact estimate[b] (SE)	90% confidence interval	*p*-Value
PY 1 through 5	15.9%	16.0%	0.2 (0.1)	(0.0, 0.4)	0.17
Percentage of index acute care hospital discharges that were followed by an unplanned acute care hospitalization or ED visit (including observation stays) within 30 days					
Baseline	25.8%	26.0%	NA	NA	NA
PY 1	25.9%	26.1%	0.1 (0.2)	(-0.2, 0.3)	0.77
PY 2	26.1%	26.2%	0.2 (0.2)	(-0.1, 0.5)	0.33
PY 3	26.1%	26.5%	-0.1 (0.2)	(-0.4, 0.2)	0.43
PY 4	25.8%	25.8%	0.3 (0.2)	(0.0, 0.6)	0.16
PY 5	25.8%	26.3%	-0.3 (0.2)	(-0.6, 0.1)	0.22
PY 1 through 5	25.9%	26.2%	0.0 (0.1)	(-0.2, 0.3)	0.90
Percentage of index ED (including observation stays) discharges that were followed by an unplanned acute care hospitalization or ED visit (including observation stays) within 30 days					
Baseline	29.5%	30.0%	NA	NA	NA
PY 1	29.2%	29.9%	-0.1 (0.2)	(-0.4, 0.2)	0.49
PY 2	29.0%	29.7%	-0.2 (0.2)	(-0.5, 0.1)	0.36
PY 3	29.0%	29.7%	-0.1 (0.2)	(-0.4, 0.2)	0.45
PY 4	29.0%	29.6%	0.0 (0.2)	(-0.4, 0.3)	0.84
PY 5	28.2%	29.1%	-0.4* (0.2)	(-0.7, 0.0)	0.10
PY 1 through 5	28.9%	29.6%	-0.2 (0.2)	(-0.4, 0.1)	0.29
Percentage of 65 and older Medicare FFS beneficiaries who received two or more prescriptions for high risk medications in the same class[d]					
Baseline	11.9%	12.1%	NA	NA	NA
PY 1	12.1%	12.3%	0.0 (0.1)	(-0.1, 0.1)	0.84
PY 2	11.9%	12.2%	-0.1* (0.1)	(-0.3, 0.0)	0.06
PY 3	14.3%	14.2%	0.2** (0.1)	(0.1, 0.4)	0.02
PY 4	14.2%	14.1%	0.2** (0.1)	(0.1, 0.4)	0.03
PY 5	13.8%	13.7%	0.2** (0.1)	(0.0, 0.4)	0.05
PY 1 through 5	13.3%	13.4%	0.1 (0.1)	(0.0, 0.2)	0.15
Long-term opioid use[i]					
Baseline	8.8%	8.6%	NA	NA	NA
PY 1	8.2%	7.9%	0.1** (0.1)	(0.0, 0.2)	0.03
PY 2	7.5%	7.2%	0.1 (0.1)	(0.0, 0.2)	0.26
PY 3	6.7%	6.5%	0.0 (0.1)	(-0.1, 0.1)	1.00
PY 4	6.1%	5.9%	-0.1 (0.1)	(-0.2, 0.1)	0.45
PY 5	5.4%	5.4%	-0.2** (0.1)	(-0.4, -0.1)	0.03
PY 1 through 5	6.7%	6.5%	0.0 (0.1)	(-0.1, 0.1)	0.99

Table 5.A.3.1a. ***(continued)***

	Track 1—Overall				
	CPC+ mean[a]	C mean[a]	Impact estimate[b] (SE)	90% confidence interval	*p*-Value
Potential opioid overuse[j]					
Baseline	19.2%	18.3%	NA	NA	NA
PY 1	17.4%	16.2%	0.3 (0.2)	(-0.1, 0.7)	0.26
PY 2	15.4%	15.0%	-0.5 (0.3)	(-0.9, 0.0)	0.11
PY 3	13.3%	13.2%	-0.8** (0.3)	(-1.3, -0.3)	0.01
PY 4	12.4%	12.5%	-0.9*** (0.3)	(-1.5, -0.4)	0.01
PY 5	11.7%	11.9%	-1.1*** (0.4)	(-1.7, -0.5)	0.00
PY 1 through 5	14.2%	13.8%	-0.4* (0.2)	(-0.9, 0.0)	0.07
Annualized number of low-value services per 1,000 beneficiaries[k]					
Baseline	366	358	NA	NA	NA
PY 1	335	327	-0.8 (1.8)	(-3.7, 2.1)	0.66
PY 2	338	329	1.0 (2.2)	(-2.6, 4.7)	0.64
PY 3	343	332	2.1 (2.5)	(-2.0, 6.3)	0.40
PY 4	291	279	3.5 (2.8)	(-1.1, 8.2)	0.22
PY 1 through 4	326	316	1.5 (2.1)	(-1.9, 4.9)	0.46
Unweighted sample sizes for other quality of care measures					
Number of index discharges for readmission	1,379,970	4,657,631			
Number of index ED discharges	2,494,107	8,720,490			
Number of 65 and older Medicare FFS beneficiaries for the high-risk medication measure	975,863	3,292,558			
Number of beneficiaries for long-term opioid use	992,417	3,392,578			
Number of beneficiaries for potential opioid overuse	96,093	314,248			
Number of beneficiaries for low-value services measure	1,261,923	4,295,094			

Source: Mathematica's analysis of Medicare claims data from January 2013 through December 2021.

Notes: For the quality-of-care outcomes, we present the absolute impact estimate only. We do so because percentage impacts for some of the binary outcomes are likely to be misleadingly large, given the low means for the outcome measures.

This table indicates which estimates are statistically significant; when we interpret evidence, we combine evidence from the magnitude of the effect, the *p*-values, findings on related outcomes, subgroups, sensitivity tests, and other data sources on model implementation.

We grouped the claims-based quality-of-care measures into separate domains according to the Comprehensive Primary Care Functions under which they appear in the 2018 CPC+ Implementation Guide (CMMI 2018).

[a] We report the actual, unadjusted averages in the baseline period which are similar for the CPC+ and comparison groups due to matching. In the intervention periods, the comparison group mean is computed by subtracting the regression adjusted difference between the CPC+ and comparison means in each time period from the CPC+ mean in that same time period.

[b] Each impact estimate is regression-adjusted using a difference-in-differences analysis that reflects the difference of the average outcome for Medicare FFS beneficiaries attributed to CPC+ practices in the five years of CPC+ to the average outcome in the baseline year, relative to the same difference over time for Medicare FFS beneficiaries attributed to comparison practices, while controlling for beneficiary characteristics and practice fixed effects.

[c] The numbers of Track 1 CPC+ and comparison practices are same as in Table 5.A.1.1a, and hence, are not reported separately in this table. The beneficiary-level measures for recommended services for diabetes, breast cancer screening, continuity of care, and opioid use are affected only by matching weights (and not by time observed) because the measures require beneficiaries to have full year of eligibility in each program year. After accounting for matching weights, the effective sample size for the comparison group for the measures presented in this table is 46 to 49 percent of the size of the actual comparison group.

Table 5.A.3.1a. ***(continued)***

[d] These measures require that beneficiaries be continuously enrolled in Medicare FFS Parts A and B as well as in Medicare Part D, and not use hospice services during the measurement year.

[e] The continuity of care measures are calculated for beneficiaries who were in the ITT sample at the beginning of the year and were FFS eligible for the full year in each program year and had qualifying ambulatory visits in the program year. Qualifying ambulatory visits are (1) office or other outpatient visit for E&M; (2) ophthalmological services; (3) medical examination and evaluation; and (4) new enrollee and annual wellness visits.

[f] For each physician, this measure indicates the percentage of beneficiaries for whom the physician was considered "most comprehensive" (i.e. saw the beneficiary for the largest share of their unique diagnoses codes) out of all beneficiaries the physician saw in the year.

[g] The new problem management measure is a score that indicates how often a primary care physician continues to treat a beneficiary's new condition versus referring the beneficiary (or the beneficiary self-referring) to a specialist or different provider. Since the new problem management measure requires a one-year look forward period, this measure is not available for PY 5 (as creating the measure for PY 5 would have required using incomplete 2022 claims data).

[h] The range of services measure is a score ranging from 0–5 that counts the number of service categories for which the physician billed. The five service categories included in the measure are: immunization, behavioral or mental health counseling, treatment of minor lacerations, cryotherapy/skin excision, and joint injection.

[i] To be included in the analysis of both long-term opioid use and potential overuse, a beneficiary had to: (1) be assigned to a practice; (2) be continuously enrolled in Medicare Parts A, B, and D throughout each calendar year or until death; and (3) have at least one opioid prescription during the measurement year. We further excluded beneficiaries for whom opioid use is appropriate: beneficiaries with a diagnosis of cancer during the measurement year or one year before, or a diagnosis of sickle cell disease or hospice use during the measurement year. The regression models for both opioid use outcomes additionally control for changes in state-level PDMP characteristics and opioid funding, and county-level opioid marketing intensity.

[j] This measure is defined only among long-term users of opioids.

[k] This measure is the annualized total number of services that provide little to no benefit to patients, have potential to cause harm, incur unnecessary costs to patients, or waste limited healthcare resources, per 1,000 beneficiaries. Because three of the low-value services are identified using a one-year look-forward period to determine whether the service was low-value or not, this measure is not available for PY 5.

*/**/*** Significantly different from zero at the 0.10/0.05/0.01 level, two-tailed test.

C = comparison; E&M = Evaluation and Management; FFS = fee-for-service; ITT = Intent-to-treat; NA = not applicable; NPI = National Provider Identifier; PDMP = prescription drug monitoring program; PY = Program Year; SE = standard error; SSP = Medicare Shared Savings Program.

Table 5.A.3.1b. Regression-adjusted means and estimated impacts of CPC+ on selected claims-based quality-of-care measures for attributed Medicare FFS beneficiaries by program year and average across the five program years, Track 1 by SSP status

	Track 1—SSP					Track 1—Non-SSP					
	CPC+ mean[a]	C mean[a]	Impact estimate[b] (SE)	90% confidence interval	*p*-Value	CPC+ mean[a]	C mean[a]	Impact estimate[b] (SE)	90% confidence interval	*p*-Value	*p*-Value for SSP vs. non-SSP difference
Planned care and population health measures for beneficiaries ages 18–75 with diabetes (percentage)											
Received HbA1c test											
Baseline	91.9%	92.1%	NA	NA	NA	89.8%	91.0%	NA	NA	NA	NA
PY 1	92.1%	92.3%	0.1 (0.2)	(-0.2, 0.4)	0.70	90.1%	91.4%	0.0 (0.2)	(-0.4, 0.4)	0.93	0.86
PY 2	92.3%	92.1%	0.4* (0.2)	(0.0, 0.8)	0.06	90.0%	91.5%	-0.2 (0.3)	(-0.6, 0.3)	0.47	0.08
PY 3	92.4%	91.8%	0.8** (0.4)	(0.2, 1.4)	0.03	90.3%	91.5%	0.0 (0.3)	(-0.5, 0.5)	0.91	0.11
PY 4	89.5%	89.0%	0.8** (0.4)	(0.2, 1.4)	0.03	87.4%	88.2%	0.5 (0.3)	(-0.1, 1.0)	0.16	0.54
PY 5	91.8%	91.2%	0.8 (0.5)	(0.0, 1.7)	0.11	89.5%	91.0%	-0.2 (0.3)	(-0.8, 0.3)	0.52	0.09
PY 1 through 5	91.6%	91.3%	0.6** (0.3)	(0.2, 1.0)	0.03	89.4%	90.7%	0.0 (0.2)	(-0.4, 0.4)	0.95	0.12
Received eye exam											
Baseline	64.6%	66.2%	NA	NA	NA	62.4%	62.6%	NA	NA	NA	NA
PY 1	65.0%	66.8%	-0.3 (0.3)	(-0.8, 0.3)	0.42	64.6%	63.1%	1.7*** (0.3)	(1.2, 2.2)	0.00	0.00
PY 2	66.2%	67.1%	0.6 (0.4)	(0.0, 1.2)	0.11	65.2%	63.4%	2.0*** (0.4)	(1.4, 2.6)	0.00	0.01
PY 3	66.2%	67.3%	0.5 (0.4)	(-0.2, 1.1)	0.25	65.1%	64.5%	0.8* (0.5)	(0.0, 1.6)	0.10	0.59
PY 4	61.4%	62.1%	0.8* (0.4)	(0.1, 1.5)	0.05	61.5%	60.2%	1.5*** (0.5)	(0.6, 2.3)	0.00	0.33
PY 5	63.7%	64.6%	0.7 (0.5)	(-0.1, 1.4)	0.17	63.5%	62.6%	1.1** (0.5)	(0.2, 2.0)	0.04	0.52
PY 1 through 5	64.5%	65.6%	0.4 (0.3)	(-0.1, 1.0)	0.19	63.9%	62.7%	1.4*** (0.4)	(0.8, 2.0)	0.00	0.05
Received attention for nephropathy											
Baseline	82.4%	81.7%	NA	NA	NA	79.3%	80.0%	NA	NA	NA	NA
PY 1	83.2%	82.0%	0.4 (0.3)	(-0.1, 0.9)	0.16	80.5%	80.3%	0.9** (0.4)	(0.3, 1.5)	0.02	0.35
PY 2	83.7%	82.1%	0.9** (0.4)	(0.3, 1.5)	0.02	81.0%	80.4%	1.3*** (0.5)	(0.5, 2.0)	0.00	0.49
PY 3	83.7%	82.7%	0.3 (0.4)	(-0.4, 0.9)	0.52	81.1%	80.5%	1.2** (0.5)	(0.4, 2.1)	0.02	0.14
PY 4	80.1%	79.2%	0.2 (0.5)	(-0.6, 1.0)	0.67	77.9%	77.3%	1.3** (0.6)	(0.2, 2.3)	0.04	0.17

Table 5.A.3.1b. ***(continued)***

	Track 1—SSP					Track 1—Non-SSP					
	CPC+ mean[a]	C mean[a]	Impact estimate[b] (SE)	90% confidence interval	*p*-Value	CPC+ mean[a]	C mean[a]	Impact estimate[b] (SE)	90% confidence interval	*p*-Value	*p*-Value for SSP vs. non-SSP difference
PY 5	82.5%	82.3%	-0.5 (0.5)	(-1.4, 0.4)	0.38	79.7%	79.7%	0.7 (0.6)	(-0.3, 1.7)	0.23	0.14
PY 1 through 5	82.6%	81.6%	0.3 (0.3)	(-0.3, 0.9)	0.39	80.0%	79.6%	1.1** (0.4)	(0.4, 1.8)	0.01	0.15
Diabetes Composite Measure 1 (received all three tests above: HbA1c test, eye exam, attention for nephropathy)											
Baseline	53.2%	53.9%	NA	NA	NA	48.8%	49.8%	NA	NA	NA	NA
PY 1	53.7%	54.8%	-0.4 (0.4)	(-1.0, 0.2)	0.32	51.1%	50.6%	1.5*** (0.4)	(0.9, 2.2)	0.00	0.00
PY 2	55.4%	55.0%	1.1*** (0.4)	(0.4, 1.8)	0.01	52.0%	50.9%	2.1*** (0.5)	(1.4, 2.9)	0.00	0.11
PY 3	55.4%	55.2%	0.9* (0.5)	(0.1, 1.7)	0.08	51.8%	51.9%	0.9* (0.6)	(0.0, 1.9)	0.10	0.96
PY 4	49.0%	48.6%	1.2** (0.5)	(0.4, 2.1)	0.02	47.0%	46.1%	1.9*** (0.6)	(0.9, 3.0)	0.00	0.38
PY 5	52.5%	52.5%	0.7 (0.6)	(-0.3, 1.7)	0.25	49.7%	49.6%	1.0* (0.6)	(0.0, 2.1)	0.10	0.68
PY 1 through 5	53.2%	53.2%	0.7* (0.4)	(0.1, 1.4)	0.08	50.3%	49.8%	1.5*** (0.4)	(0.8, 2.2)	0.00	0.18
Diabetes Composite Measure 2 (received none of the three tests above)											
Baseline	2.3%	2.1%	NA	NA	NA	2.7%	2.5%	NA	NA	NA	NA
PY 1	2.1%	2.1%	-0.2** (0.1)	(-0.4, -0.1)	0.03	2.5%	2.4%	-0.2 (0.1)	(-0.4, 0.0)	0.14	0.79
PY 2	2.2%	2.1%	-0.1 (0.1)	(-0.3, 0.1)	0.28	2.4%	2.4%	-0.3** (0.1)	(-0.5, -0.1)	0.02	0.28
PY 3	2.1%	2.1%	-0.1 (0.1)	(-0.3, 0.1)	0.48	2.4%	2.4%	-0.3* (0.1)	(-0.5, 0.0)	0.05	0.27
PY 4	3.4%	3.3%	-0.1 (0.1)	(-0.3, 0.1)	0.47	3.6%	3.7%	-0.3* (0.2)	(-0.6, 0.0)	0.07	0.32
PY 5	2.3%	2.3%	-0.1 (0.1)	(-0.3, 0.1)	0.36	2.7%	2.6%	-0.1 (0.1)	(-0.3, 0.1)	0.51	0.88
PY 1 through 5	2.4%	2.4%	-0.1 (0.1)	(-0.3, 0.0)	0.16	2.7%	2.7%	-0.2** (0.1)	(-0.4, 0.0)	0.04	0.46
Unweighted sample sizes for the diabetes measures[c]											
Number of beneficiaries	135,434	530,386				131,192	393,924				
Number of beneficiary-years	398,579	1,550,483				391,318	1,150,082				

Table 5.A.3.1b. ***(continued)***

	Track 1—SSP					Track 1—Non-SSP					
	CPC+ mean[a]	C mean[a]	Impact estimate[b] (SE)	90% confidence interval	*p*-Value	CPC+ mean[a]	C mean[a]	Impact estimate[b] (SE)	90% confidence interval	*p*-Value	*p*-Value for SSP vs. non-SSP difference
Planned care and population health measures for female beneficiaries ages 52–74 (percentage)											
Received breast cancer screening											
Baseline	73.6%	74.0%	NA	NA	NA	71.5%	72.3%	NA	NA	NA	NA
PY 1	74.3%	74.6%	0.1 (0.2)	(-0.2, 0.4)	0.67	72.7%	72.7%	0.7*** (0.2)	(0.4, 1.1)	0.00	0.03
PY 2	74.9%	75.2%	0.2 (0.3)	(-0.3, 0.6)	0.52	73.8%	72.8%	1.7*** (0.3)	(1.2, 2.2)	0.00	0.00
PY 3	75.4%	75.8%	0.0 (0.3)	(-0.5, 0.5)	0.92	74.4%	73.6%	1.6*** (0.3)	(1.1, 2.1)	0.00	0.00
PY 4	73.1%	73.7%	-0.2 (0.3)	(-0.8, 0.3)	0.52	72.9%	72.1%	1.5*** (0.4)	(0.9, 2.1)	0.00	0.00
PY 5	73.3%	74.0%	-0.2 (0.4)	(-0.9, 0.4)	0.54	73.2%	72.1%	1.9*** (0.4)	(1.2, 2.5)	0.00	0.00
PY 1 through 5	74.2%	74.6%	0.0 (0.3)	(-0.4, 0.4)	0.94	73.4%	72.7%	1.5*** (0.3)	(1.0, 2.0)	0.00	0.00
Unweighted sample sizes for the breast cancer screening measure[c]											
Number of beneficiaries	225,441	868,239				215,559	622,802				
Number of beneficiary-years	688,653	2,641,313				668,706	1,905,248				
Planned care and population health measures for beneficiaries ages 21 and older[d]											
Percentage of beneficiaries with cardiovascular disease who were prescribed and filled statin therapy											
Baseline	58.6%	59.6%	NA	NA	NA	59.2%	58.5%	NA	NA	NA	NA
PY 1	60.0%	61.0%	-0.1 (0.1)	(-0.4, 0.1)	0.27	60.5%	59.7%	0.1 (0.1)	(-0.1, 0.4)	0.43	0.19
PY 2	58.9%	60.2%	-0.3** (0.2)	(-0.6, -0.1)	0.04	60.0%	59.2%	0.0 (0.2)	(-0.3, 0.4)	0.94	0.17
PY 3	60.3%	61.5%	-0.2 (0.2)	(-0.6, 0.1)	0.22	61.1%	60.4%	-0.1 (0.2)	(-0.5, 0.3)	0.72	0.62
PY 4	61.1%	62.3%	-0.3 (0.2)	(-0.7, 0.0)	0.16	61.7%	61.2%	-0.2 (0.3)	(-0.7, 0.2)	0.42	0.79
PY 5	62.8%	64.2%	-0.4* (0.3)	(-0.9, 0.0)	0.09	63.2%	62.6%	-0.1 (0.3)	(-0.5, 0.3)	0.70	0.36
PY 1 through 5	60.7%	61.9%	-0.3* (0.2)	(-0.6, 0.0)	0.08	61.3%	60.7%	-0.1 (0.2)	(-0.4, 0.3)	0.77	0.37
Unweighted sample sizes for the statin therapy measure[c]											
Number of beneficiaries	446,546	1,659,996				391,559	1,155,940				
Number of beneficiary-years	1,494,423	5,520,443				1,321,635	3,827,236				

Table 5.A.3.1b. ***(continued)***

	Track 1—SSP					Track 1—Non-SSP					
	CPC+ mean[a]	C mean[a]	Impact estimate[b] (SE)	90% confidence interval	*p*-Value	CPC+ mean[a]	C mean[a]	Impact estimate[b] (SE)	90% confidence interval	*p*-Value	*p*-Value for SSP vs. non-SSP difference
Planned care and population health measures for beneficiaries ages 18 and older[d]											
Percentage of beneficiaries on diabetes medications with proportion of days covered by medication > 80%											
Baseline	77.7%	77.9%	NA	NA	NA	76.5%	76.5%	NA	NA	NA	NA
PY 1	78.7%	78.9%	-0.1 (0.3)	(-0.6, 0.4)	0.85	77.6%	78.2%	-0.6* (0.4)	(-1.2, -0.1)	0.07	0.21
PY 2	79.6%	79.7%	0.1 (0.3)	(-0.5, 0.6)	0.86	78.8%	79.3%	-0.6 (0.4)	(-1.2, 0.0)	0.11	0.18
PY 3	79.1%	79.4%	-0.2 (0.3)	(-0.7, 0.4)	0.64	78.2%	78.7%	-0.6 (0.4)	(-1.2, 0.1)	0.16	0.43
PY 4	82.3%	82.2%	0.2 (0.3)	(-0.4, 0.8)	0.53	81.8%	81.7%	0.1 (0.4)	(-0.6, 0.7)	0.83	0.81
PY 5	81.2%	81.6%	-0.3 (0.3)	(-0.8, 0.3)	0.47	80.6%	81.2%	-0.7* (0.4)	(-1.3, 0.0)	0.10	0.42
PY 1 through 5	80.3%	80.4%	0.0 (0.3)	(-0.5, 0.4)	0.88	79.5%	79.9%	-0.5 (0.3)	(-1.0, 0.0)	0.12	0.27
Percentage of beneficiaries on renin-angiotensin system antagonists with proportion of days covered by medication > 80%											
Baseline	79.3%	79.0%	NA	NA	NA	78.4%	78.1%	NA	NA	NA	NA
PY 1	81.3%	81.6%	-0.6*** (0.2)	(-0.9, -0.3)	0.00	80.7%	80.4%	0.1 (0.2)	(-0.2, 0.4)	0.67	0.02
PY 2	82.2%	82.3%	-0.3* (0.2)	(-0.6, 0.0)	0.10	81.7%	81.3%	0.1 (0.2)	(-0.3, 0.4)	0.68	0.16
PY 3	82.1%	82.5%	-0.7*** (0.2)	(-1.0, -0.3)	0.00	81.6%	81.8%	-0.4** (0.2)	(-0.8, -0.1)	0.04	0.42
PY 4	84.3%	84.5%	-0.4** (0.2)	(-0.8, -0.1)	0.03	83.8%	83.9%	-0.4* (0.2)	(-0.8, -0.1)	0.05	0.90
PY 5	83.9%	84.5%	-0.8*** (0.2)	(-1.1, -0.4)	0.00	83.5%	83.8%	-0.6*** (0.2)	(-1.0, -0.3)	0.00	0.63
PY 1 through 5	82.8%	83.2%	-0.6*** (0.2)	(-0.8, -0.3)	0.00	82.3%	82.4%	-0.3 (0.2)	(-0.6, 0.0)	0.11	0.23
Percentage of beneficiaries on statins with proportion of days covered by medication > 80%											
Baseline	76.2%	76.8%	NA	NA	NA	75.8%	75.7%	NA	NA	NA	NA
PY 1	76.1%	76.9%	-0.2 (0.2)	(-0.6, 0.1)	0.21	75.8%	75.5%	0.2 (0.2)	(-0.2, 0.5)	0.45	0.16
PY 2	79.5%	80.0%	0.0 (0.2)	(-0.3, 0.2)	0.79	79.4%	79.2%	0.1 (0.2)	(-0.2, 0.5)	0.50	0.49
PY 3	79.6%	80.4%	-0.3 (0.2)	(-0.6, 0.0)	0.15	79.4%	79.1%	0.2 (0.2)	(-0.1, 0.6)	0.28	0.08
PY 4	82.5%	83.0%	0.0 (0.2)	(-0.3, 0.3)	0.98	82.5%	82.6%	-0.2 (0.3)	(-0.6, 0.3)	0.55	0.63
PY 5	81.9%	82.7%	-0.4* (0.2)	(-0.7, 0.0)	0.09	81.9%	81.9%	-0.1 (0.2)	(-0.4, 0.3)	0.81	0.33

Table 5.A.3.1b. ***(continued)***

	Track 1—SSP					Track 1—Non-SSP					
	CPC+ mean[a]	C mean[a]	Impact estimate[b] (SE)	90% confidence interval	*p*-Value	CPC+ mean[a]	C mean[a]	Impact estimate[b] (SE)	90% confidence interval	*p*-Value	*p*-Value for SSP vs. non-SSP difference
PY 1 through 5	80.1%	80.8%	-0.2 (0.2)	(-0.4, 0.1)	0.24	80.0%	79.9%	0.1 (0.2)	(-0.2, 0.4)	0.70	0.29
Percentage of beneficiaries with both coronary artery disease (CAD) and diabetes who were prescribed and filled angiotensin-converting enzyme (ACE) inhibitor or angiotensin receptor blocker (ARB) therapy											
Baseline	78.4%	77.5%	NA	NA	NA	78.0%	77.9%	NA	NA	NA	NA
PY 1	78.6%	77.4%	0.3 (0.3)	(-0.2, 0.9)	0.35	77.9%	77.8%	0.0 (0.4)	(-0.6, 0.7)	0.91	0.61
PY 2	77.9%	76.7%	0.4 (0.4)	(-0.2, 1.0)	0.32	77.6%	77.7%	-0.2 (0.5)	(-1.0, 0.6)	0.62	0.32
PY 3	77.7%	76.8%	0.1 (0.4)	(-0.6, 0.7)	0.85	77.9%	77.8%	0.1 (0.5)	(-0.7, 0.9)	0.84	0.96
PY 4	77.1%	76.1%	0.1 (0.4)	(-0.6, 0.8)	0.79	77.0%	76.5%	0.4 (0.5)	(-0.5, 1.3)	0.48	0.69
PY 5	77.0%	76.2%	-0.1 (0.4)	(-0.9, 0.6)	0.76	77.6%	76.8%	0.7 (0.5)	(-0.2, 1.6)	0.19	0.23
PY 1 through 5	77.6%	76.6%	0.2 (0.3)	(-0.4, 0.7)	0.61	77.6%	77.3%	0.2 (0.4)	(-0.5, 0.9)	0.66	0.96
Unweighted sample sizes for percentage of beneficiaries on diabetes medications with proportion of days covered by medication > 80%											
Number of beneficiaries	98,176	373,966				90,865	268,554				
Number of beneficiary-years	293,197	1,112,488				273,121	795,850				
Unweighted sample sizes for percentage of beneficiaries on renin-angiotensin system antagonists with proportion of days covered by medication > 80%											
Number of beneficiaries	292,384	1,120,504				273,783	799,121				
Number of beneficiary-years	907,992	3,484,182				867,803	2,472,190				
Unweighted sample sizes for percentage of beneficiaries on statins with proportion of days covered by medication > 80%											
Number of beneficiaries	342,319	1,313,752				310,343	908,444				
Number of beneficiary-years	1,108,170	4,268,031				1,020,597	2,929,376				
Unweighted sample sizes for percentage of beneficiaries with both CAD and diabetes who were prescribed and filled ACE inhibitor or ARB therapy											
Number of beneficiaries	85,836	293,162				72,976	210,877				
Number of beneficiary-years	188,573	644,594				162,557	458,688				
Measures for continuity of care[d]											
Percentage of primary care ambulatory visits at assigned practice											
Baseline	75.7%	74.1%	NA	NA	NA	75.3%	73.3%	NA	NA	NA	NA
PY 1	72.5%	71.1%	-0.1 (0.3)	(-0.7, 0.4)	0.68	72.4%	70.1%	0.3 (0.4)	(-0.3, 0.9)	0.48	0.43
PY 2	63.8%	62.2%	0.0 (0.7)	(-1.1, 1.1)	0.97	64.1%	61.2%	0.8 (0.6)	(-0.3, 1.8)	0.21	0.41
PY 3	61.2%	59.0%	0.7 (0.7)	(-0.5, 1.9)	0.37	61.6%	58.5%	1.0 (0.7)	(-0.2, 2.2)	0.16	0.73
PY 4	53.5%	52.2%	-0.3 (1.0)	(-2.0, 1.3)	0.75	55.5%	51.1%	2.3* (1.2)	(0.3, 4.4)	0.05	0.09

Table 5.A.3.1b. ***(continued)***

	Track 1—SSP					Track 1—Non-SSP					
	CPC+ mean[a]	C mean[a]	Impact estimate[b] (SE)	90% confidence interval	*p*-Value	CPC+ mean[a]	C mean[a]	Impact estimate[b] (SE)	90% confidence interval	*p*-Value	*p*-Value for SSP vs. non-SSP difference
PY 5	49.5%	48.3%	-0.4 (1.0)	(-2.1, 1.3)	0.69	51.4%	48.1%	1.2 (1.1)	(-0.6, 2.9)	0.28	0.29
PY 1 through 5	59.7%	58.1%	0.0 (0.6)	(-1.0, 1.0)	0.99	60.5%	57.3%	1.1* (0.7)	(0.0, 2.2)	0.10	0.23
Across all PCPs and specialists providing care to a patient, where each practitioner in the beneficiary's assigned practice is treated separately											
Percentage of visits with the usual provider of care (UPC)											
Baseline	47.5%	47.8%	NA	NA	NA	49.3%	49.0%	NA	NA	NA	NA
PY 1	46.5%	46.8%	0.0 (0.1)	(-0.1, 0.2)	0.86	48.3%	48.1%	-0.2 (0.1)	(-0.4, 0.0)	0.12	0.18
PY 2	45.3%	45.7%	-0.1 (0.1)	(-0.2, 0.1)	0.57	47.1%	47.0%	-0.2* (0.1)	(-0.5, 0.0)	0.07	0.30
PY 3	44.7%	44.9%	0.1 (0.1)	(-0.1, 0.3)	0.40	46.3%	46.2%	-0.3** (0.1)	(-0.5, -0.1)	0.04	0.04
PY 4	47.0%	47.6%	-0.3** (0.1)	(-0.5, -0.1)	0.04	48.5%	48.7%	-0.5*** (0.2)	(-0.8, -0.2)	0.00	0.25
PY 5	44.2%	44.6%	-0.1 (0.1)	(-0.3, 0.1)	0.59	45.7%	45.9%	-0.5*** (0.2)	(-0.7, -0.2)	0.01	0.08
PY 1 through 5	45.5%	45.9%	-0.1 (0.1)	(-0.2, 0.1)	0.58	47.1%	47.1%	-0.3*** (0.1)	(-0.5, -0.1)	0.01	0.08
Reversed Bice-Boxerman fragmentation of care index											
Baseline	77.6%	77.7%	NA	NA	NA	76.2%	76.7%	NA	NA	NA	NA
PY 1	78.5%	78.7%	-0.1 (0.1)	(-0.2, 0.1)	0.44	77.3%	77.6%	0.1 (0.1)	(-0.1, 0.3)	0.34	0.22
PY 2	79.6%	79.8%	0.0 (0.1)	(-0.2, 0.2)	0.95	78.5%	78.8%	0.1 (0.1)	(-0.1, 0.4)	0.34	0.48
PY 3	80.3%	80.6%	-0.2 (0.1)	(-0.3, 0.0)	0.20	79.3%	79.6%	0.1 (0.1)	(-0.1, 0.3)	0.47	0.18
PY 4	80.6%	80.8%	0.0 (0.1)	(-0.3, 0.2)	0.80	79.6%	79.8%	0.2 (0.2)	(-0.1, 0.5)	0.27	0.31
PY 5	81.8%	81.8%	0.1 (0.1)	(-0.2, 0.3)	0.67	80.8%	81.0%	0.3 (0.2)	(0.0, 0.6)	0.14	0.34
PY 1 through 5	80.2%	80.4%	0.0 (0.1)	(-0.2, 0.1)	0.67	79.2%	79.4%	0.2 (0.1)	(-0.1, 0.4)	0.22	0.21
Across all PCPs and specialists providing care to a patient, where all practitioners in the beneficiary's assigned practice are treated as a single practitioner											
Percentage of visits with the usual provider of care											
Baseline	49.9%	50.2%	NA	NA	NA	52.1%	51.9%	NA	NA	NA	NA
PY 1	48.8%	49.2%	-0.1 (0.1)	(-0.2, 0.1)	0.51	50.9%	50.9%	-0.1 (0.1)	(-0.3, 0.0)	0.20	0.57
PY 2	47.1%	47.6%	-0.2** (0.1)	(-0.4, 0.0)	0.04	49.2%	49.1%	-0.2 (0.2)	(-0.4, 0.1)	0.30	0.70

Table 5.A.3.1b. ***(continued)***

	Track 1—SSP					Track 1—Non-SSP					
	CPC+ mean[a]	C mean[a]	Impact estimate[b] (SE)	90% confidence interval	*p*-Value	CPC+ mean[a]	C mean[a]	Impact estimate[b] (SE)	90% confidence interval	*p*-Value	*p*-Value for SSP vs. non-SSP difference
PY 3	47.0%	47.3%	0.0 (0.1)	(-0.3, 0.2)	0.85	49.1%	49.0%	-0.1 (0.2)	(-0.4, 0.2)	0.49	0.68
PY 4	48.8%	49.5%	-0.4** (0.2)	(-0.7, -0.1)	0.01	50.7%	50.6%	-0.1 (0.2)	(-0.5, 0.2)	0.57	0.31
PY 5	45.9%	46.3%	-0.2 (0.1)	(-0.4, 0.1)	0.28	47.9%	47.9%	-0.2 (0.2)	(-0.6, 0.2)	0.42	0.91
PY 1 through 5	47.5%	47.9%	-0.2* (0.1)	(-0.3, 0.0)	0.09	49.5%	49.5%	-0.1 (0.1)	(-0.4, 0.1)	0.33	0.88
Reversed Bice-Boxerman fragmentation of care index											
Baseline	75.0%	75.1%	NA	NA	NA	73.1%	73.5%	NA	NA	NA	NA
PY 1	76.0%	76.2%	0.0 (0.1)	(-0.1, 0.2)	0.86	74.3%	74.6%	0.1 (0.1)	(-0.1, 0.3)	0.59	0.75
PY 2	77.8%	77.8%	0.2 (0.1)	(0.0, 0.4)	0.11	76.2%	76.4%	0.1 (0.2)	(-0.2, 0.4)	0.61	0.62
PY 3	77.9%	78.1%	0.0 (0.2)	(-0.3, 0.2)	0.86	76.2%	76.6%	0.0 (0.2)	(-0.4, 0.3)	0.83	0.95
PY 4	78.7%	78.8%	0.1 (0.2)	(-0.2, 0.4)	0.52	77.1%	77.7%	-0.2 (0.3)	(-0.6, 0.3)	0.55	0.39
PY 5	80.1%	80.1%	0.1 (0.2)	(-0.1, 0.4)	0.36	78.5%	78.8%	0.0 (0.3)	(-0.4, 0.4)	0.95	0.69
PY 1 through 5	78.2%	78.2%	0.1 (0.1)	(-0.1, 0.3)	0.45	76.5%	76.9%	0.0 (0.2)	(-0.3, 0.3)	0.99	0.67
Unweighted sample sizes for percentage of primary care ambulatory visits at assigned practice[c]											
Number of beneficiaries	684,621	2,671,277				646,747	1,901,696				
Number of beneficiary-years	2,378,753	9,262,260				2,293,628	6,585,570				
Unweighted sample sizes for percentage of visits with the usual provider of care[c]											
Number of beneficiaries	701,076	2,735,035				660,460	1,946,847				
Number of beneficiary-years	2,516,869	9,779,144				2,415,671	6,949,735				
Unweighted sample sizes for reversed Bice-Boxerman fragmentation of care index[c]											
Number of beneficiaries	641,558	2,486,963				599,718	1,758,304				
Number of beneficiary-years	2,088,583	8,028,496				1,967,606	5,627,835				
Comprehensiveness of care (measured at the physician level)											
Physician involvement in patient conditions											
Baseline	66.1%	67.5%	NA	NA	NA	69.7%	68.8%	NA	NA	NA	NA
PY 1	66.2%	67.3%	0.4 (0.3)	(-0.2, 0.9)	0.29	70.0%	68.6%	0.5 (0.4)	(-0.1, 1.2)	0.18	0.77
PY 2	66.4%	67.6%	0.2 (0.4)	(-0.4, 0.9)	0.59	70.4%	68.4%	1.1** (0.5)	(0.3, 1.9)	0.02	0.14
PY 3	67.5%	68.3%	0.6 (0.4)	(-0.1, 1.4)	0.15	70.9%	69.2%	0.8 (0.6)	(-0.1, 1.8)	0.13	0.78

Table 5.A.3.1b. ***(continued)***

	Track 1—SSP					Track 1—Non-SSP					
	CPC+ mean[a]	C mean[a]	Impact estimate[b] (SE)	90% confidence interval	*p*-Value	CPC+ mean[a]	C mean[a]	Impact estimate[b] (SE)	90% confidence interval	*p*-Value	*p*-Value for SSP vs. non-SSP difference
PY 4	69.9%	70.4%	0.9* (0.5)	(0.1, 1.8)	0.07	73.4%	71.8%	0.7 (0.6)	(-0.2, 1.7)	0.18	0.80
PY 5	67.5%	68.4%	0.5 (0.6)	(-0.4, 1.5)	0.34	70.9%	69.4%	0.7 (0.6)	(-0.3, 1.7)	0.24	0.85
PY 1 through 5	67.4%	68.3%	0.5 (0.3)	(-0.1, 1.1)	0.15	71.0%	69.4%	0.8** (0.4)	(0.2, 1.4)	0.04	0.57
Range of services provided by physicians											
Baseline	1.98	1.96	NA	NA	NA	2.18	1.90	NA	NA	NA	NA
PY 1	1.93	1.91	-0.01 (0.02)	(-0.05, 0.03)	0.69	2.14	1.86	0.01 (0.03)	(-0.03, 0.05)	0.72	0.59
PY 2	1.89	1.86	0.01 (0.03)	(-0.04, 0.05)	0.78	2.11	1.82	0.02 (0.03)	(-0.03, 0.07)	0.59	0.82
PY 3	1.87	1.85	0.00 (0.03)	(-0.05, 0.05)	0.96	2.04	1.82	-0.05 (0.03)	(-0.11, 0.00)	0.10	0.23
PY 4	1.76	1.78	-0.04 (0.03)	(-0.09, 0.01)	0.23	1.97	1.78	-0.08** (0.04)	(-0.14, -0.01)	0.04	0.41
PY 5	1.66	1.71	-0.07* (0.04)	(-0.13, 0.00)	0.09	1.94	1.67	-0.01 (0.04)	(-0.07, 0.06)	0.85	0.29
PY 1 through 5	1.84	1.83	-0.01 (0.02)	(-0.05, 0.02)	0.48	2.05	1.79	-0.02 (0.03)	(-0.06, 0.02)	0.50	0.95
Management of new problems by physicians											
Baseline	0.996	1.002	NA	NA	NA	1.007	1.003	NA	NA	NA	NA
PY 1	0.995	1.000	0.001 (0.002)	(-0.003, 0.005)	0.74	1.008	1.002	0.001 (0.003)	(-0.004, 0.005)	0.78	1.00
PY 2	0.995	1.001	0.000 (0.003)	(-0.004, 0.004)	0.99	1.005	0.998	0.002 (0.003)	(-0.003, 0.008)	0.51	0.60
PY 3	0.995	0.999	0.002 (0.003)	(-0.002, 0.007)	0.45	1.004	1.003	-0.003 (0.004)	(-0.009, 0.004)	0.49	0.32
PY 4	0.994	0.999	0.001 (0.004)	(-0.005, 0.007)	0.82	1.006	1.003	-0.002 (0.004)	(-0.008, 0.004)	0.65	0.63
PY 1 through 4	0.995	1.000	0.001 (0.002)	(-0.003, 0.004)	0.68	1.006	1.001	0.000 (0.003)	(-0.004, 0.004)	0.99	0.79
Unweighted sample sizes for physician involvement in patient conditions											
Number of physicians	2,198	9,113				2,014	6,921				
Number of physician-years	10,398	42,859				9,631	31,140				
Unweighted sample sizes for range of services provided by physicians											
Number of physicians	2,136	8,812				1,960	6,492				
Number of physician-years	10,077	40,816				9,369	28,884				
Unweighted sample sizes for management of new problems by physicians											
Number of physicians	2,152	8,916				1,986	6,731				
Number of physician-years	8,853	36,343				8,196	26,311				

Table 5.A.3.1b. ***(continued)***

	Track 1—SSP					Track 1—Non-SSP					
	CPC+ mean[a]	C mean[a]	Impact estimate[b] (SE)	90% confidence interval	*p*-Value	CPC+ mean[a]	C mean[a]	Impact estimate[b] (SE)	90% confidence interval	*p*-Value	*p*-Value for SSP vs. non-SSP difference
Patient and caregiver engagement											
Percentage of beneficiaries receiving hospice services											
Baseline	2.7%	2.7%	NA	NA	NA	2.8%	2.7%	NA	NA	NA	NA
PY 1	2.7%	2.7%	0.1** (0.0)	(0.0, 0.2)	0.01	2.8%	2.6%	0.0 (0.0)	(-0.1, 0.1)	0.88	0.14
PY 2	2.9%	2.8%	0.1*** (0.0)	(0.1, 0.2)	0.00	2.9%	2.8%	0.0 (0.1)	(-0.1, 0.1)	0.98	0.02
PY 3	3.1%	3.0%	0.2*** (0.0)	(0.1, 0.2)	0.00	3.1%	2.9%	0.1 (0.1)	(0.0, 0.2)	0.11	0.27
PY 4	3.3%	3.2%	0.1*** (0.0)	(0.1, 0.2)	0.01	3.3%	3.1%	0.1 (0.1)	(0.0, 0.2)	0.34	0.32
PY 5	3.3%	3.1%	0.2*** (0.1)	(0.1, 0.3)	0.00	3.3%	3.1%	0.1 (0.1)	(0.0, 0.2)	0.23	0.13
PY 1 through 5	3.1%	3.0%	0.1*** (0.0)	(0.1, 0.2)	0.00	3.1%	2.9%	0.0 (0.0)	(0.0, 0.1)	0.30	0.07
Length of hospice stay, in days (for beneficiaries receiving hospice services)											
Baseline	60	65	NA	NA	NA	60	66	NA	NA	NA	NA
PY 1	62	66	1.2 (1.4)	(-1.2, 3.5)	0.41	62	66	2.1 (1.4)	(-0.3, 4.4)	0.14	0.65
PY 2	65	68	1.7 (1.4)	(-0.6, 4.1)	0.23	68	70	3.8** (1.7)	(1.0, 6.5)	0.03	0.36
PY 3	71	72	3.1** (1.5)	(0.7, 5.5)	0.03	72	74	3.2* (1.7)	(0.4, 5.9)	0.06	0.97
PY 4	68	69	3.9** (1.6)	(1.2, 6.5)	0.02	70	72	3.2* (1.9)	(0.1, 6.2)	0.09	0.78
PY 5	70	71	4.2*** (1.6)	(1.6, 6.8)	0.01	70	72	4.0** (1.7)	(1.2, 6.8)	0.02	0.93
PY 1 through 5	68	70	2.8** (1.2)	(0.8, 4.9)	0.02	69	71	3.3** (1.4)	(1.0, 5.5)	0.02	0.81
Length of hospice stay, in days (for all beneficiaries)											
Baseline	1.6	1.8	NA	NA	NA	1.7	1.8	NA	NA	NA	NA
PY 1	1.7	1.8	0.1** (0.0)	(0.0, 0.2)	0.02	1.7	1.7	0.1 (0.0)	(0.0, 0.1)	0.18	0.57
PY 2	1.9	1.9	0.1*** (0.1)	(0.1, 0.2)	0.00	2.0	2.0	0.1* (0.1)	(0.0, 0.2)	0.06	0.60
PY 3	2.2	2.1	0.2*** (0.1)	(0.1, 0.3)	0.00	2.2	2.2	0.2** (0.1)	(0.1, 0.3)	0.01	0.43
PY 4	2.3	2.2	0.2*** (0.1)	(0.1, 0.4)	0.00	2.3	2.2	0.1** (0.1)	(0.0, 0.3)	0.03	0.25
PY 5	2.3	2.2	0.3*** (0.1)	(0.2, 0.4)	0.00	2.3	2.2	0.2** (0.1)	(0.1, 0.3)	0.01	0.21

Table 5.A.3.1b. ***(continued)***

	Track 1—SSP					Track 1—Non-SSP					
	CPC+ mean[a]	C mean[a]	Impact estimate[b] (SE)	90% confidence interval	*p*-Value	CPC+ mean[a]	C mean[a]	Impact estimate[b] (SE)	90% confidence interval	*p*-Value	*p*-Value for SSP vs. non-SSP difference
PY 1 through 5	2.1	2.0	0.2*** (0.0)	(0.1, 0.3)	0.00	2.1	2.1	0.1** (0.1)	(0.0, 0.2)	0.01	0.31
Unweighted sample sizes for patient and caregiver engagement measures											
Number of beneficiaries for length of hospice stay	72,266	264,758				69,708	191,252				
Other quality measures											
Percentage of index acute care hospital discharges that were followed by an unplanned readmission within 30 days											
Baseline	15.4%	15.8%	NA	NA	NA	15.6%	15.7%	NA	NA	NA	NA
PY 1	15.3%	15.7%	0.1 (0.2)	(-0.3, 0.4)	0.72	16.0%	15.9%	0.2 (0.2)	(-0.2, 0.5)	0.42	0.73
PY 2	15.9%	15.9%	0.4* (0.2)	(0.0, 0.7)	0.06	15.8%	15.8%	0.1 (0.2)	(-0.3, 0.5)	0.74	0.32
PY 3	15.9%	16.1%	0.2 (0.2)	(-0.2, 0.5)	0.37	15.8%	16.0%	-0.2 (0.2)	(-0.5, 0.2)	0.48	0.26
PY 4	16.2%	16.2%	0.5** (0.2)	(0.1, 0.9)	0.02	16.3%	16.0%	0.3 (0.2)	(-0.1, 0.7)	0.19	0.52
PY 5	16.2%	16.4%	0.2 (0.2)	(-0.1, 0.6)	0.28	16.0%	16.1%	0.0 (0.2)	(-0.4, 0.4)	0.93	0.51
PY 1 through 5	15.9%	16.1%	0.3 (0.2)	(0.0, 0.6)	0.11	16.0%	16.0%	0.1 (0.2)	(-0.2, 0.4)	0.69	0.42
Percentage of index acute care hospital discharges that were followed by an unplanned acute care hospitalization or ED visit (including observation stays) within 30 days											
Baseline	25.3%	25.7%	NA	NA	NA	26.2%	26.3%	NA	NA	NA	NA
PY 1	25.1%	25.8%	-0.3 (0.2)	(-0.7, 0.1)	0.30	26.7%	26.4%	0.4 (0.3)	(-0.1, 0.8)	0.15	0.08
PY 2	25.8%	26.0%	0.3 (0.2)	(-0.1, 0.7)	0.28	26.4%	26.4%	0.1 (0.3)	(-0.4, 0.5)	0.76	0.61
PY 3	25.9%	26.3%	0.0 (0.2)	(-0.4, 0.4)	0.90	26.3%	26.7%	-0.3 (0.3)	(-0.8, 0.1)	0.23	0.32
PY 4	25.4%	25.6%	0.2 (0.3)	(-0.2, 0.7)	0.37	26.2%	25.9%	0.3 (0.3)	(-0.1, 0.8)	0.24	0.81
PY 5	25.6%	26.2%	-0.2 (0.3)	(-0.7, 0.2)	0.42	26.1%	26.4%	-0.2 (0.3)	(-0.7, 0.2)	0.41	0.97
PY 1 through 5	25.6%	26.0%	0.0 (0.2)	(-0.3, 0.3)	0.93	26.3%	26.4%	0.0 (0.2)	(-0.3, 0.4)	0.89	0.96
Percentage of index ED (including observation stays) discharges that were followed by an unplanned acute care hospitalization or ED visit (including observation stays) within 30 days											
Baseline	28.6%	29.2%	NA	NA	NA	30.3%	30.8%	NA	NA	NA	NA
PY 1	28.5%	29.2%	-0.2 (0.2)	(-0.5, 0.2)	0.50	30.0%	30.6%	-0.1 (0.3)	(-0.5, 0.3)	0.75	0.84
PY 2	28.1%	28.8%	-0.1 (0.2)	(-0.5, 0.3)	0.74	29.8%	30.6%	-0.2 (0.3)	(-0.7, 0.2)	0.35	0.65
PY 3	28.2%	28.9%	-0.1 (0.2)	(-0.5, 0.3)	0.56	29.8%	30.4%	-0.1 (0.3)	(-0.6, 0.3)	0.61	0.98

Table 5.A.3.1b. ***(continued)***

	Track 1—SSP					Track 1—Non-SSP					
	CPC+ mean[a]	C mean[a]	Impact estimate[b] (SE)	90% confidence interval	*p*-Value	CPC+ mean[a]	C mean[a]	Impact estimate[b] (SE)	90% confidence interval	*p*-Value	*p*-Value for SSP vs. non-SSP difference
PY 4	28.6%	28.9%	0.3 (0.3)	(-0.2, 0.7)	0.35	29.5%	30.4%	-0.3 (0.3)	(-0.9, 0.3)	0.37	0.20
PY 5	27.8%	28.6%	-0.3 (0.3)	(-0.8, 0.2)	0.38	28.6%	29.7%	-0.5 (0.3)	(-1.0, 0.0)	0.12	0.62
PY 1 through 5	28.2%	28.9%	-0.1 (0.2)	(-0.4, 0.3)	0.68	29.6%	30.3%	-0.2 (0.2)	(-0.6, 0.1)	0.29	0.61
Percentage of 65 and older Medicare FFS beneficiaries who received two or more prescriptions for high risk medications in the same class[d]											
Baseline	11.6%	11.6%	NA	NA	NA	12.1%	12.5%	NA	NA	NA	NA
PY 1	11.8%	11.9%	-0.1 (0.1)	(-0.2, 0.0)	0.27	12.5%	12.8%	0.1 (0.1)	(-0.1, 0.2)	0.42	0.18
PY 2	11.5%	11.8%	-0.2** (0.1)	(-0.4, -0.1)	0.01	12.3%	12.7%	0.0 (0.1)	(-0.2, 0.2)	0.88	0.12
PY 3	14.0%	13.9%	0.1 (0.1)	(-0.1, 0.4)	0.32	14.6%	14.6%	0.4** (0.2)	(0.1, 0.6)	0.02	0.24
PY 4	13.8%	13.7%	0.1 (0.1)	(-0.1, 0.4)	0.29	14.5%	14.6%	0.3** (0.2)	(0.1, 0.6)	0.03	0.36
PY 5	13.4%	13.3%	0.2 (0.1)	(-0.1, 0.4)	0.23	14.1%	14.2%	0.2 (0.2)	(0.0, 0.5)	0.14	0.78
PY 1 through 5	13.0%	13.0%	0.0 (0.1)	(-0.1, 0.2)	0.84	13.7%	13.8%	0.2* (0.1)	(0.0, 0.4)	0.08	0.24
Long-term opioid use											
Baseline	8.4%	7.8%	NA	NA	NA	9.2%	9.4%	NA	NA	NA	NA
PY 1	7.8%	7.3%	0.0 (0.1)	(-0.2, 0.1)	0.51	8.6%	8.5%	0.3*** (0.1)	(0.2, 0.4)	0.00	0.00
PY 2	7.1%	6.6%	-0.2* (0.1)	(-0.3, 0.0)	0.08	7.9%	7.7%	0.4*** (0.1)	(0.2, 0.6)	0.00	0.00
PY 3	6.3%	6.1%	-0.3*** (0.1)	(-0.5, -0.1)	0.00	7.1%	6.9%	0.4*** (0.1)	(0.2, 0.6)	0.00	0.00
PY 4	5.8%	5.6%	-0.4*** (0.1)	(-0.6, -0.2)	0.00	6.4%	6.3%	0.3** (0.1)	(0.0, 0.5)	0.05	0.00
PY 5	5.1%	5.0%	-0.5*** (0.1)	(-0.7, -0.3)	0.00	5.7%	5.8%	0.1 (0.1)	(-0.1, 0.3)	0.53	0.00
PY 1 through 5	6.4%	6.0%	-0.2*** (0.1)	(-0.4, -0.1)	0.01	7.1%	7.0%	0.3*** (0.1)	(0.1, 0.5)	0.00	0.00
Potential opioid overuse											
Baseline	19.9%	18.9%	NA	NA	NA	18.4%	17.8%	NA	NA	NA	NA
PY 1	18.2%	17.2%	0.0 (0.3)	(-0.5, 0.5)	0.92	16.5%	15.3%	0.5 (0.3)	(-0.1, 1.1)	0.15	0.30
PY 2	16.2%	16.5%	-1.4*** (0.4)	(-2.1, -0.8)	0.00	14.7%	13.6%	0.4 (0.4)	(-0.3, 1.2)	0.32	0.00
PY 3	14.5%	14.5%	-1.0** (0.4)	(-1.8, -0.3)	0.02	12.1%	12.0%	-0.6 (0.5)	(-1.4, 0.2)	0.20	0.48

Table 5.A.3.1b. ***(continued)***

	Track 1—SSP					Track 1—Non-SSP					
	CPC+ mean[a]	C mean[a]	Impact estimate[b] (SE)	90% confidence interval	*p*-Value	CPC+ mean[a]	C mean[a]	Impact estimate[b] (SE)	90% confidence interval	*p*-Value	*p*-Value for SSP vs. non-SSP difference
PY 4	13.8%	14.0%	-1.2** (0.5)	(-2.0, -0.4)	0.01	11.1%	11.2%	-0.8 (0.5)	(-1.6, 0.1)	0.13	0.52
PY 5	13.0%	13.2%	-1.2** (0.5)	(-2.1, -0.3)	0.02	10.4%	10.9%	-1.1** (0.5)	(-2.0, -0.3)	0.03	0.95
PY 1 through 5	15.3%	15.0%	-0.8** (0.3)	(-1.3, -0.2)	0.02	13.1%	12.6%	-0.1 (0.4)	(-0.8, 0.5)	0.71	0.21
Annualized number of low-value services per 1,000 beneficiaries											
Baseline	387	371	NA	NA	NA	345	344	NA	NA	NA	NA
PY 1	352	340	-4.5* (2.6)	(-8.8, -0.3)	0.08	317	312	3.3 (2.4)	(-0.6, 7.2)	0.17	0.03
PY 2	354	340	-1.6 (2.9)	(-6.5, 3.2)	0.58	321	316	3.9 (3.3)	(-1.6, 9.3)	0.24	0.21
PY 3	361	344	1.0 (3.3)	(-4.5, 6.5)	0.77	324	319	3.4 (3.8)	(-2.8, 9.7)	0.37	0.63
PY 4	306	288	2.4 (4.1)	(-4.3, 9.1)	0.55	276	270	5.2 (3.9)	(-1.1, 11.6)	0.18	0.62
PY 1 through 4	343	327	-0.7 (2.8)	(-5.3, 4.0)	0.82	308	303	3.9 (2.9)	(-0.9, 8.8)	0.18	0.26
Unweighted sample sizes for unplanned readmission, receiving hospice services, and length of hospice stay[g]											
Number of index discharges for readmission	708,023	2,722,429				671,947	1,935,202				
Number of index ED discharges	1,230,885	4,899,507				1,263,222	3,820,983				
Number of 65 and older Medicare FFS beneficiaries for the high risk medication measure	510,344	1,947,695				466,868	1,352,663				
Number of beneficiaries for long-term opioid use	512,852	1,985,920				480,933	1,414,950				
Number of beneficiaries for potential opioid overuse	46,908	172,431				49,292	142,465				
Number of beneficiaries for low-value services measure	651,417	2,507,748				612,582	1,799,451				

Source: Mathematica's analysis of Medicare claims data from January 2013 through December 2021.

Notes: For the quality-of-care outcomes, we present the absolute impact estimate only. We do so because percentage impacts for some of the binary outcomes are likely to be misleadingly large, given the low means for the outcome measures.

This table indicates which estimates are statistically significant; when we interpret evidence, we combine evidence from the magnitude of the effect, the *p*-values, findings on related outcomes, subgroups, sensitivity tests, and other data sources on model implementation.

We grouped the claims-based quality-of-care measures into separate domains according to the Comprehensive Primary Care Functions under which they appear in the 2018 CPC+ Implementation Guide (CMMI 2018).

[a] We report the actual, unadjusted averages in the baseline period which are similar for the CPC+ and comparison groups due to matching. In the intervention periods, the comparison group mean is computed by subtracting the regression adjusted difference between the CPC+ and comparison means in each time period from the CPC+ mean in that same time period.

Table 5.A.3.1b. ***(continued)***

[b] Each impact estimate is regression-adjusted using a difference-in-differences analysis that reflects the difference of the average outcome for Medicare FFS beneficiaries attributed to CPC+ practices in the five years of CPC+ to the average outcome in the baseline year, relative to the same difference over time for Medicare FFS beneficiaries attributed to comparison practices, while controlling for beneficiary characteristics and practice fixed effects.

[c] The numbers of Track 1 CPC+ and comparison practices are same as in Table 5.A.1.1b and hence, are not reported separately in this table. The beneficiary-level measures for recommended services for diabetes, breast cancer screening, continuity of care, and opioid use are affected only by matching weights (and not by time observed) because the measures require beneficiaries to have full year of eligibility in each program year. After accounting for matching weights, the effective sample size for the comparison group for the measures presented in this table is 43 to 54 percent of the size of the actual comparison group.

[d] These measures require that beneficiaries be continuously enrolled in Medicare FFS Parts A and B as well as in Medicare Part D, and not use hospice services during the measurement year.

[e] The continuity of care measures are calculated for beneficiaries who were in the ITT sample at the beginning of the year and were FFS eligible for the full year in each program year and had qualifying ambulatory visits in the program year. Qualifying ambulatory visits are (1) office or other outpatient visit for E&M; (2) ophthalmological services; (3) medical examination and evaluation; and (4) new enrollee and annual wellness visits.

[f] For each physician, this measure indicates the percentage of beneficiaries for whom the physician was considered “most comprehensive” (i.e. saw the beneficiary for the largest share of their unique diagnoses codes) out of all beneficiaries the physician saw in the year.

[g] The new problem management measure is a score that indicates how often a primary care physician continues to treat a beneficiary’s new condition versus referring the beneficiary (or the beneficiary self-referring) to a specialist or different provider. Since the new problem management measure requires a one-year look forward period, this measure is not available for PY 5 (as creating the measure for PY 5 would have required using incomplete 2022 claims data).

[h] The range of services measure is a score ranging from 0–5 that counts the number of service categories for which the physician billed. The five service categories included in the measure are: immunization, behavioral or mental health counseling, treatment of minor lacerations, cryotherapy/skin excision, and joint injection.

[i] To be included in the analysis of both long-term opioid use and potential overuse, a beneficiary had to: (1) be assigned to a practice; (2) be continuously enrolled in Medicare Parts A, B, and D throughout each calendar year or until death; and (3) have at least one opioid prescription during the measurement year. We further excluded beneficiaries for whom opioid use is appropriate: beneficiaries with a diagnosis of cancer during the measurement year or one year before, or a diagnosis of sickle cell disease or hospice use during the measurement year. The regression models for both opioid use outcomes additionally control for changes in state-level PDMP characteristics and opioid funding, and county-level opioid marketing intensity.

[j] This measure is defined only among long-term users of opioids.

[k] This measure is the annualized total number of services that provide little to no benefit to patients, have potential to cause harm, incur unnecessary costs to patients, or waste limited healthcare resources, per 1,000 beneficiaries. Because three of the low-value services are identified using a one-year look-forward period to determine whether the service was low-value or not, this measure is not available for PY 5.

*/**/*** Significantly different from zero at the 0.10/0.05/0.01 level, two-tailed test.

C = comparison; E&M = Evaluation and Management; FFS = fee-for-service; ITT = Intent-to-treat; NA = not applicable; NPI = National Provider Identifier; PDMP = prescription drug monitoring program; PY = Program Year; SE = standard error; SSP = Medicare Shared Savings Program.

Table 5.A.3.2a. Regression-adjusted means and estimated impacts of CPC+ on selected claims-based quality-of-care measures for attributed Medicare FFS beneficiaries by program year and average across the five program years, Track 2

	Track 2—Overall				
	CPC+ mean[a]	C mean[a]	Impact estimate[b] (SE)	90% confidence interval	*p*-Value
Planned care and population health measures for beneficiaries ages 18–75 with diabetes (percentage)					
Received HbA1c test					
Baseline	92.5%	92.1%	NA	NA	NA
PY 1	92.9%	92.3%	0.3 (0.2)	(0.0, 0.6)	0.12
PY 2	92.7%	92.1%	0.2 (0.2)	(-0.1, 0.5)	0.23
PY 3	92.7%	92.0%	0.4 (0.2)	(0.0, 0.7)	0.12
PY 4	89.9%	89.3%	0.3 (0.2)	(-0.1, 0.7)	0.26
PY 5	92.2%	91.4%	0.4 (0.3)	(0.0, 0.9)	0.14
PY 1 through 5	92.0%	91.4%	0.3* (0.2)	(0.0, 0.6)	0.10
Received eye exam					
Baseline	65.4%	65.5%	NA	NA	NA
PY 1	66.1%	66.3%	-0.1 (0.2)	(-0.4, 0.3)	0.83
PY 2	67.3%	66.3%	1.2*** (0.3)	(0.7, 1.7)	0.00
PY 3	67.9%	66.6%	1.5*** (0.3)	(0.9, 2.0)	0.00
PY 4	63.4%	62.3%	1.3*** (0.4)	(0.7, 1.9)	0.00
PY 5	65.8%	64.6%	1.4*** (0.4)	(0.8, 2.0)	0.00
PY 1 through 5	66.1%	65.2%	1.1*** (0.3)	(0.6, 1.5)	0.00
Received attention for nephropathy					
Baseline	82.7%	82.1%	NA	NA	NA
PY 1	83.4%	82.6%	0.3 (0.2)	(-0.1, 0.6)	0.26
PY 2	84.0%	82.7%	0.8*** (0.3)	(0.3, 1.2)	0.01
PY 3	83.8%	82.9%	0.4 (0.3)	(-0.1, 0.9)	0.23
PY 4	80.0%	80.1%	-0.6* (0.4)	(-1.2, -0.1)	0.07
PY 5	82.5%	82.0%	0.0 (0.3)	(-0.6, 0.5)	0.93
PY 1 through 5	82.7%	82.0%	0.2 (0.3)	(-0.3, 0.6)	0.51
Diabetes Composite Measure 1 (received all three tests above: HbA1c test, eye exam, attention for nephropathy)					
Baseline	54.0%	53.6%	NA	NA	NA
PY 1	55.1%	54.6%	0.1 (0.3)	(-0.3, 0.6)	0.63
PY 2	56.5%	54.5%	1.6*** (0.3)	(1.0, 2.2)	0.00
PY 3	56.7%	55.0%	1.4*** (0.4)	(0.7, 2.0)	0.00
PY 4	50.7%	49.5%	0.8* (0.4)	(0.1, 1.5)	0.05
PY 5	54.1%	52.5%	1.3*** (0.4)	(0.6, 2.0)	0.00
PY 1 through 5	54.6%	53.2%	1.1*** (0.3)	(0.5, 1.6)	0.00

Table 5.A.3.2a. ***(continued)***

	Track 2—Overall				
	CPC+ mean[a]	C mean[a]	Impact estimate[b] (SE)	90% confidence interval	*p*-Value
Diabetes Composite Measure 2 (received none of the three tests above)					
Baseline	2.1%	2.1%	NA	NA	NA
PY 1	1.9%	2.0%	-0.1 (0.1)	(-0.2, 0.1)	0.33
PY 2	2.0%	2.1%	-0.2* (0.1)	(-0.3, 0.0)	0.06
PY 3	1.9%	2.2%	-0.2*** (0.1)	(-0.4, -0.1)	0.01
PY 4	3.3%	3.2%	0.1 (0.1)	(-0.1, 0.2)	0.41
PY 5	2.2%	2.3%	-0.1 (0.1)	(-0.2, 0.1)	0.36
PY 1 through 5	2.3%	2.4%	-0.1 (0.1)	(-0.2, 0.0)	0.16
Unweighted sample sizes for the diabetes measures[c]					
Number of beneficiaries	325,766	763,565			
Number of beneficiary-years	964,361	2,243,525			
Planned care and population health measures for female beneficiaries ages 52–74 (percentage)					
Received breast cancer screening					
Baseline	73.6%	74.3%	NA	NA	NA
PY 1	74.7%	74.9%	0.5*** (0.1)	(0.2, 0.7)	0.00
PY 2	75.4%	75.2%	0.9*** (0.2)	(0.5, 1.2)	0.00
PY 3	76.0%	75.7%	0.9*** (0.2)	(0.6, 1.3)	0.00
PY 4	73.9%	73.8%	0.8*** (0.3)	(0.4, 1.2)	0.00
PY 5	74.0%	74.0%	0.7** (0.3)	(0.2, 1.1)	0.01
PY 1 through 5	74.8%	74.7%	0.8*** (0.2)	(0.4, 1.1)	0.00
Unweighted sample sizes for the breast cancer screening measure[c]					
Number of beneficiaries	535,011	1,245,621			
Number of beneficiary-years	1,648,354	3,816,802			
Planned care and population health measures for beneficiaries ages 21 and older[d]					
Percentage of beneficiaries with cardiovascular disease who were prescribed and filled statin therapy					
Baseline	59.4%	59.7%	NA	NA	NA
PY 1	60.7%	60.9%	0.0 (0.1)	(-0.2, 0.2)	0.90
PY 2	59.7%	60.0%	-0.1 (0.1)	(-0.3, 0.2)	0.65
PY 3	60.9%	61.3%	-0.1 (0.2)	(-0.4, 0.2)	0.50
PY 4	61.5%	62.0%	-0.2 (0.2)	(-0.5, 0.1)	0.18
PY 5	62.9%	63.4%	-0.3 (0.2)	(-0.6, 0.0)	0.14
PY 1 through 5	61.2%	61.6%	-0.1 (0.1)	(-0.4, 0.1)	0.33

Table 5.A.3.2a. ***(continued)***

	Track 2—Overall				
	CPC+ mean[a]	C mean[a]	Impact estimate[b] (SE)	90% confidence interval	*p*-Value
Unweighted sample sizes for the statin therapy measure[c]					
Number of beneficiaries	991,346	2,356,827			
Number of beneficiary-years	3,328,316	7,874,418			
Planned care and population health measures for beneficiaries ages 18 and older[d]					
Percentage of beneficiaries on diabetes medications with proportion of days covered by medication > 80%					
Baseline	77.4%	77.5%	NA	NA	NA
PY 1	78.6%	78.7%	0.0 (0.3)	(-0.4, 0.4)	0.95
PY 2	79.5%	79.5%	0.1 (0.3)	(-0.3, 0.6)	0.59
PY 3	78.7%	79.1%	-0.2 (0.3)	(-0.6, 0.2)	0.44
PY 4	82.2%	82.0%	0.3 (0.3)	(-0.1, 0.7)	0.23
PY 5	81.1%	81.4%	-0.2 (0.3)	(-0.6, 0.2)	0.41
PY 1 through 5	80.1%	80.2%	0.0 (0.2)	(-0.3, 0.4)	0.99
Percentage of beneficiaries on renin-angiotensin system antagonists with proportion of days covered by medication > 80%					
Baseline	78.7%	78.8%	NA	NA	NA
PY 1	81.1%	81.2%	0.1 (0.1)	(-0.1, 0.3)	0.50
PY 2	82.0%	82.1%	0.0 (0.1)	(-0.2, 0.3)	0.78
PY 3	81.9%	82.4%	-0.3** (0.2)	(-0.6, -0.1)	0.02
PY 4	84.0%	84.5%	-0.3** (0.2)	(-0.6, -0.1)	0.04
PY 5	83.6%	84.4%	-0.7*** (0.2)	(-0.9, -0.4)	0.00
PY 1 through 5	82.6%	83.0%	-0.2* (0.1)	(-0.5, 0.0)	0.05
Percentage of beneficiaries on statins with proportion of days covered by medication > 80%					
Baseline	76.4%	76.5%	NA	NA	NA
PY 1	76.6%	76.4%	0.2 (0.1)	(0.0, 0.5)	0.12
PY 2	79.8%	79.8%	0.1 (0.1)	(-0.2, 0.3)	0.61
PY 3	79.8%	80.1%	-0.2 (0.1)	(-0.4, 0.1)	0.24
PY 4	82.8%	83.1%	-0.3* (0.2)	(-0.5, 0.0)	0.07
PY 5	82.1%	82.6%	-0.4*** (0.1)	(-0.7, -0.2)	0.00
PY 1 through 5	80.4%	80.6%	-0.1 (0.1)	(-0.3, 0.1)	0.30
Percentage of beneficiaries with both coronary artery disease (CAD) and diabetes who were prescribed and filled angiotensin-converting enzyme (ACE) inhibitor or angiotensin receptor blocker (ARB) therapy					
Baseline	78.5%	78.1%	NA	NA	NA
PY 1	78.3%	77.9%	0.0 (0.3)	(-0.4, 0.5)	0.93
PY 2	77.7%	77.1%	0.2 (0.3)	(-0.3, 0.7)	0.54
PY 3	77.7%	77.2%	0.1 (0.3)	(-0.4, 0.7)	0.64

Table 5.A.3.2a. ***(continued)***

	Track 2—Overall				
	CPC+ mean[a]	C mean[a]	Impact estimate[b] (SE)	90% confidence interval	*p*-Value
PY 4	76.4%	76.1%	-0.1 (0.3)	(-0.6, 0.5)	0.83
PY 5	76.8%	76.3%	0.0 (0.3)	(-0.5, 0.6)	0.93
PY 1 through 5	77.4%	76.9%	0.1 (0.3)	(-0.3, 0.5)	0.79
Unweighted sample sizes for percentage of beneficiaries on diabetes medications with proportion of days covered by medication> 80%					
Number of beneficiaries	224,764	531,906			
Number of beneficiary-years	676,013	1,588,630			
Unweighted sample sizes for percentage of beneficiaries on renin-angiotensin system antagonists with proportion of days covered by medication > 80%					
Number of beneficiaries	678,940	1,611,531			
Number of beneficiary-years	2,134,990	5,039,973			
Unweighted sample sizes for percentage of beneficiaries on statins with proportion of days covered by medication > 80%					
Number of beneficiaries	781,199	1,876,281			
Number of beneficiary-years	2,559,902	6,124,360			
Unweighted sample sizes for percentage of beneficiaries with both CAD and diabetes who were prescribed and filled ACE inhibitor or ARB therapy					
Number of beneficiaries	179,419	414,972			
Number of beneficiary-years	403,782	916,725			
Measures for continuity of care[e]					
Percentage of primary care ambulatory visits at assigned practice					
Baseline	75.5%	73.3%	NA	NA	NA
PY 1	72.9%	70.4%	0.3 (0.2)	(-0.1, 0.8)	0.16
PY 2	64.7%	61.6%	1.0** (0.4)	(0.3, 1.7)	0.02
PY 3	62.7%	58.9%	1.6*** (0.5)	(0.9, 2.4)	0.00
PY 4	55.7%	52.5%	1.0 (0.8)	(-0.2, 2.3)	0.17
PY 5	51.3%	48.3%	0.8 (0.7)	(-0.3, 2.0)	0.24
PY 1 through 5	61.0%	57.8%	1.0** (0.4)	(0.3, 1.7)	0.02
Across all PCPs and specialists providing care to a patient, where each practitioner in the beneficiary's assigned practice is treated separately					
Percentage of visits with the usual provider of care (UPC)					
Baseline	48.0%	47.9%	NA	NA	NA
PY 1	47.0%	47.0%	0.0 (0.1)	(-0.2, 0.1)	0.55
PY 2	46.0%	46.0%	-0.1 (0.1)	(-0.2, 0.0)	0.16
PY 3	45.2%	45.2%	-0.1 (0.1)	(-0.3, 0.0)	0.15
PY 4	47.7%	47.8%	-0.3** (0.1)	(-0.4, -0.1)	0.03

Table 5.A.3.2a. ***(continued)***

	Track 2—Overall				
	CPC+ mean[a]	C mean[a]	Impact estimate[b] (SE)	90% confidence interval	*p*-Value
PY 5	44.9%	45.1%	-0.3** (0.1)	(-0.5, -0.1)	0.01
PY 1 through 5	46.1%	46.2%	-0.2** (0.1)	(-0.3, 0.0)	0.04
Reversed Bice-Boxerman fragmentation of care index					
Baseline	77.7%	77.9%	NA	NA	NA
PY 1	78.6%	78.8%	0.1 (0.1)	(0.0, 0.2)	0.30
PY 2	79.7%	79.8%	0.1* (0.1)	(0.0, 0.3)	0.09
PY 3	80.5%	80.6%	0.1 (0.1)	(0.0, 0.3)	0.15
PY 4	80.7%	80.8%	0.1 (0.1)	(-0.1, 0.3)	0.31
PY 5	81.8%	81.8%	0.3** (0.1)	(0.1, 0.5)	0.01
PY 1 through 5	80.3%	80.4%	0.2* (0.1)	(0.0, 0.3)	0.07
Across all PCPs and specialists providing care to a patient, where all practitioners in the beneficiary's assigned practice are treated as a single practitioner					
Percentage of visits with the usual provider of care (UPC)					
Baseline	51.3%	51.1%	NA	NA	NA
PY 1	50.3%	50.0%	0.1 (0.1)	(-0.1, 0.2)	0.46
PY 2	48.6%	48.4%	-0.1 (0.1)	(-0.2, 0.1)	0.59
PY 3	48.6%	48.3%	0.0 (0.1)	(-0.2, 0.2)	0.83
PY 4	50.3%	50.2%	-0.2 (0.2)	(-0.5, 0.0)	0.18
PY 5	47.4%	47.4%	-0.2 (0.2)	(-0.5, 0.0)	0.17
PY 1 through 5	49.0%	48.8%	-0.1 (0.1)	(-0.2, 0.1)	0.47
Reversed Bice-Boxerman fragmentation of care index					
Baseline	74.1%	74.5%	NA	NA	NA
PY 1	75.1%	75.6%	-0.1 (0.1)	(-0.2, 0.1)	0.52
PY 2	76.9%	77.2%	0.1 (0.1)	(-0.1, 0.2)	0.50
PY 3	76.9%	77.3%	0.0 (0.1)	(-0.3, 0.2)	0.75
PY 4	77.9%	78.1%	0.1 (0.2)	(-0.2, 0.4)	0.57
PY 5	79.2%	79.4%	0.2 (0.2)	(0.0, 0.5)	0.17
PY 1 through 5	77.3%	77.6%	0.1 (0.1)	(-0.1, 0.3)	0.60
Unweighted sample sizes for percentage of primary care ambulatory visits at assigned practice[c]					
Number of beneficiaries	1,619,135	3,853,209			
Number of beneficiary-years	5,717,632	13,479,179			
Unweighted sample sizes for percentage of visits with the usual provider of care[c]					
Number of beneficiaries	1,650,218	3,937,174			
Number of beneficiary-years	5,988,153	14,182,812			

Table 5.A.3.2a. ***(continued)***

	Track 2—Overall				
	CPC+ mean[a]	C mean[a]	Impact estimate[b] (SE)	90% confidence interval	*p*-Value
Unweighted sample sizes for reversed Bice-Boxerman fragmentation of care index[c]					
Number of beneficiaries	1,494,458	3,569,815			
Number of beneficiary-years	4,879,711	11,549,075			
Comprehensiveness of care (measured at the physician level)					
Physician involvement in patient conditions[f]					
Baseline	67.1%	67.3%	NA	NA	NA
PY 1	66.9%	67.2%	-0.1 (0.3)	(-0.6, 0.3)	0.68
PY 2	66.9%	67.2%	-0.1 (0.3)	(-0.6, 0.4)	0.78
PY 3	67.9%	68.1%	-0.1 (0.4)	(-0.7, 0.5)	0.84
PY 4	70.5%	70.6%	0.0 (0.4)	(-0.6, 0.7)	0.93
PY 5	68.2%	68.3%	0.1 (0.4)	(-0.6, 0.8)	0.88
PY 1 through 5	68.0%	68.2%	-0.1 (0.3)	(-0.5, 0.4)	0.84
Range of services provided by physicians[h]					
Baseline	2.04	1.92	NA	NA	NA
PY 1	1.99	1.88	0.00 (0.02)	(-0.03, 0.03)	0.97
PY 2	1.98	1.83	0.03 (0.02)	(0.00, 0.06)	0.15
PY 3	1.98	1.82	0.04* (0.02)	(0.00, 0.08)	0.08
PY 4	1.89	1.77	0.00 (0.03)	(-0.05, 0.05)	1.00
PY 5	1.84	1.71	0.01 (0.03)	(-0.04, 0.05)	0.80
PY 1 through 5	1.94	1.81	0.02 (0.02)	(-0.01, 0.05)	0.35
Management of new problems by physicians[g]					
Baseline	0.997	0.997	NA	NA	NA
PY 1	0.995	0.998	-0.004* (0.002)	(-0.007, 0.000)	0.09
PY 2	0.995	0.997	-0.002 (0.002)	(-0.006, 0.001)	0.31
PY 3	0.995	0.997	-0.003 (0.002)	(-0.007, 0.001)	0.25
PY 4	0.993	0.997	-0.005* (0.003)	(-0.009, 0.000)	0.08
PY 1 through 4	0.995	0.997	-0.003* (0.002)	(-0.006, 0.000)	0.08
Unweighted sample sizes for Physician involvement in patient conditions					
Number of physicians	5,726	13,748			
Number of physician-years	27,650	63,539			
Unweighted sample sizes for Range of services provided by physicians[g=]					
Number of physicians	5,521	13,140			
Number of physician-years	26,739	60,015			

Table 5.A.3.2a. ***(continued)***

	Track 2—Overall				
	CPC+ mean[a]	C mean[a]	Impact estimate[b] (SE)	90% confidence interval	*p*-Value
Unweighted sample sizes for Management of new problems by physicians					
Number of physicians	5,595	13,434			
Number of physician-years	23,374	53,794			
Patient and caregiver engagement					
Percentage of beneficiaries receiving hospice services					
Baseline	2.8%	2.7%	NA	NA	NA
PY 1	2.8%	2.7%	0.0 (0.0)	(0.0, 0.1)	0.12
PY 2	3.0%	2.8%	0.1*** (0.0)	(0.1, 0.2)	0.00
PY 3	3.1%	2.9%	0.2*** (0.0)	(0.1, 0.2)	0.00
PY 4	3.3%	3.2%	0.1*** (0.0)	(0.0, 0.2)	0.00
PY 5	3.3%	3.1%	0.2*** (0.0)	(0.1, 0.2)	0.00
PY 1 through 5	3.1%	3.0%	0.1*** (0.0)	(0.1, 0.2)	0.00
Length of hospice stay, in days (for beneficiaries receiving hospice services)					
Baseline	62	67	NA	NA	NA
PY 1	62	66	0.5 (1.0)	(-1.2, 2.2)	0.60
PY 2	66	69	2.4** (1.1)	(0.6, 4.3)	0.03
PY 3	71	72	3.9*** (1.2)	(2.0, 5.9)	0.00
PY 4	69	72	2.3* (1.2)	(0.3, 4.2)	0.06
PY 5	70	73	2.1* (1.2)	(0.2, 4.1)	0.07
PY 1 through 5	68	71	2.3** (0.9)	(0.8, 3.9)	0.01
Length of hospice stay, in days (for all beneficiaries)					
Baseline	1.7	1.8	NA	NA	NA
PY 1	1.7	1.8	0.0 (0.0)	(0.0, 0.1)	0.15
PY 2	2.0	1.9	0.2*** (0.0)	(0.1, 0.2)	0.00
PY 3	2.2	2.1	0.2*** (0.0)	(0.2, 0.3)	0.00
PY 4	2.3	2.3	0.2*** (0.0)	(0.1, 0.2)	0.00
PY 5	2.3	2.2	0.2*** (0.1)	(0.1, 0.3)	0.00
PY 1 through 5	2.1	2.1	0.2*** (0.0)	(0.1, 0.2)	0.00
Unweighted sample sizes for patient and caregiver engagement measures					
Number of beneficiaries for length of hospice stay	174,905	390,437			

Table 5.A.3.2a. ***(continued)***

	Track 2—Overall				
	CPC+ mean[a]	C mean[a]	Impact estimate[b] (SE)	90% confidence interval	*p*-Value
Other quality measures					
Percentage of index acute care hospital discharges that were followed by an unplanned readmission within 30 days					
Baseline	15.6%	15.8%	NA	NA	NA
PY 1	15.6%	15.9%	-0.1 (0.2)	(-0.3, 0.2)	0.70
PY 2	15.9%	16.0%	0.0 (0.2)	(-0.2, 0.3)	0.85
PY 3	15.9%	16.2%	-0.1 (0.2)	(-0.4, 0.2)	0.56
PY 4	16.1%	16.1%	0.2 (0.2)	(-0.1, 0.4)	0.29
PY 5	16.3%	16.1%	0.3** (0.2)	(0.1, 0.6)	0.05
PY 1 through 5	16.0%	16.1%	0.1 (0.1)	(-0.2, 0.3)	0.64
Percentage of index acute care hospital discharges that were followed by an unplanned acute care hospitalization or ED visit (including observation stays) within 30 days					
Baseline	25.8%	26.1%	NA	NA	NA
PY 1	25.9%	26.2%	0.0 (0.2)	(-0.3, 0.2)	0.78
PY 2	26.1%	26.3%	0.0 (0.2)	(-0.3, 0.3)	0.95
PY 3	26.2%	26.5%	-0.1 (0.2)	(-0.4, 0.2)	0.69
PY 4	25.6%	25.8%	0.0 (0.2)	(-0.3, 0.3)	0.96
PY 5	25.9%	26.1%	0.0 (0.2)	(-0.3, 0.3)	0.82
PY 1 through 5	25.9%	26.2%	0.0 (0.1)	(-0.3, 0.2)	0.87
Percentage of index ED (including observation stays) discharges that were followed by an unplanned acute care hospitalization or ED visit (including observation stays) within 30 days					
Baseline	29.3%	29.7%	NA	NA	NA
PY 1	29.0%	29.4%	0.0 (0.2)	(-0.3, 0.3)	1.00
PY 2	29.1%	29.5%	0.0 (0.2)	(-0.3, 0.3)	0.94
PY 3	29.3%	29.6%	0.1 (0.2)	(-0.2, 0.4)	0.67
PY 4	29.6%	29.7%	0.4* (0.2)	(0.0, 0.7)	0.09
PY 5	29.0%	29.1%	0.4 (0.2)	(0.0, 0.7)	0.11
PY 1 through 5	29.2%	29.5%	0.1 (0.2)	(-0.1, 0.4)	0.39
Percentage of 65 and older Medicare FFS beneficiaries who received two or more prescriptions for high risk medications in the same class[d]					
Baseline	11.9%	11.8%	NA	NA	NA
PY 1	12.1%	12.1%	0.0 (0.1)	(-0.1, 0.1)	0.55
PY 2	12.0%	11.9%	-0.1 (0.1)	(-0.2, 0.1)	0.38
PY 3	14.1%	14.0%	0.0 (0.1)	(-0.1, 0.2)	0.60
PY 4	14.0%	13.8%	0.1 (0.1)	(-0.1, 0.2)	0.31
PY 5	13.6%	13.5%	0.1 (0.1)	(-0.1, 0.2)	0.49
PY 1 through 5	13.2%	13.1%	0.0 (0.1)	(-0.1, 0.1)	0.74

Table 5.A.3.2a. ***(continued)***

	Track 2—Overall				
	CPC+ mean[a]	C mean[a]	Impact estimate[b] (SE)	90% confidence interval	*p*-Value
Long-term opioid use[j]					
Baseline	8.9%	8.5%	NA	NA	NA
PY 1	8.2%	7.7%	0.1 (0.1)	(0.0, 0.2)	0.20
PY 2	7.5%	7.1%	0.0 (0.1)	(-0.1, 0.2)	0.54
PY 3	6.6%	6.3%	-0.1 (0.1)	(-0.2, 0.1)	0.48
PY 4	6.0%	5.7%	-0.1 (0.1)	(-0.2, 0.1)	0.32
PY 5	5.4%	5.1%	-0.1 (0.1)	(-0.3, 0.0)	0.12
PY 1 through 5	6.7%	6.3%	0.0 (0.1)	(-0.1, 0.1)	0.75
Potential opioid overuse[j]					
Baseline	19.5%	19.2%	NA	NA	NA
PY 1	17.7%	17.4%	0.1 (0.2)	(-0.3, 0.5)	0.77
PY 2	15.5%	15.7%	-0.4 (0.3)	(-0.9, 0.1)	0.19
PY 3	13.5%	14.0%	-0.7** (0.3)	(-1.3, -0.2)	0.04
PY 4	12.5%	13.0%	-0.7* (0.4)	(-1.3, -0.1)	0.05
PY 5	11.8%	12.3%	-0.8* (0.4)	(-1.4, -0.1)	0.05
PY 1 through 5	14.3%	14.5%	-0.4 (0.3)	(-0.8, 0.0)	0.13
Annualized number of low-value services per 1,000 beneficiaries[k]					
Baseline	344	346	NA	NA	NA
PY 1	314	316	-0.4 (1.8)	(-3.3, 2.5)	0.81
PY 2	315	318	-1.4 (2.1)	(-4.8, 2.1)	0.51
PY 3	319	322	-0.8 (2.6)	(-5.1, 3.5)	0.77
PY 4	270	273	-1.0 (2.7)	(-5.5, 3.4)	0.70
PY 1 through 4	304	306	-0.9 (2.0)	(-4.3, 2.4)	0.65
Unweighted sample sizes for other quality of care measures					
Number of index discharges for readmission	1,696,607	3,957,516			
Number of index ED discharges	3,034,910	7,240,593			
Number of 65 and older Medicare FFS beneficiaries for the high-risk medication measure	1,180,840	2,796,228			
Number of beneficiaries for long-term opioid use	1,205,132	2,861,935			
Number of beneficiaries for potential opioid overuse	116,871	255,657			

Table 5.A.3.2a. ***(continued)***

	Track 2—Overall				
	CPC+ mean[a]	C mean[a]	Impact estimate[b] (SE)	90% confidence interval	*p*-Value
Number of beneficiaries for low-value services measure	1,527,374	3,625,150			

Source: Mathematica's analysis of Medicare claims data from January 2013 through December 2021.

Notes: For the quality-of-care outcomes, we present the absolute impact estimate only. We do so because percentage impacts for some of the binary outcomes are likely to be misleadingly large, given the low means for the outcome measures.

This table indicates which estimates are statistically significant; when we interpret evidence, we combine evidence from the magnitude of the effect, the *p*-values, findings on related outcomes, subgroups, sensitivity tests, and other data sources on model implementation.

We grouped the claims-based quality-of-care measures into separate domains according to the Comprehensive Primary Care Functions under which they appear in the 2018 CPC+ Implementation Guide (CMMI 2018).

[a] We report the actual, unadjusted averages in the baseline period which are similar for the CPC+ and comparison groups due to matching. In the intervention periods, the comparison group mean is computed by subtracting the regression adjusted difference between the CPC+ and comparison means in each time period from the CPC+ mean in that same time period.

[b] Each impact estimate is regression-adjusted using a difference-in-differences analysis that reflects the difference of the average outcome for Medicare FFS beneficiaries attributed to CPC+ practices in the five years of CPC+ to the average outcome in the baseline year, relative to the same difference over time for Medicare FFS beneficiaries attributed to comparison practices, while controlling for beneficiary characteristics and practice fixed effects.

[c] The numbers of Track 2 CPC+ and comparison practices are same as in Tables 5.A.1.2a, and hence, are not reported separately in this table. The beneficiary-level measures for recommended services for diabetes, breast cancer screening, continuity of care, and opioid use are affected only by matching weights (and not by time observed) because the measures require beneficiaries to have full year of eligibility in each program year. After accounting for matching weights, the effective sample size for the comparison group for the measures presented in this table is 40 to 42 percent of the size of the actual comparison group.

[d] These measures require that beneficiaries be continuously enrolled in Medicare FFS Parts A and B as well as in Medicare Part D, and not use hospice services during the measurement year.

[e] The continuity of care measures are calculated for beneficiaries who were in the ITT sample at the beginning of the year and were FFS eligible for the full year in each program year and had qualifying ambulatory visits in the program year. Qualifying ambulatory visits are (1) office or other outpatient visit for E&M; (2) ophthalmological services; (3) medical examination and evaluation; and (4) new enrollee and annual wellness visits.

[f] For each physician, this measure indicates the percentage of beneficiaries for whom the physician was considered "most comprehensive" (i.e. saw the beneficiary for the largest share of their unique diagnoses codes) out of all beneficiaries the physician saw in the year.

[g] The new problem management measure is a score that indicates how often a primary care physician continues to treat a beneficiary's new condition versus referring the beneficiary (or the beneficiary self-referring) to a specialist or different provider. Since the new problem management measure requires a one-year look forward period, this measure is not available for PY 5 (as creating the measure for PY 5 would have required using incomplete 2022 claims data).

[h] The range of services measure is a score ranging from 0–5 that counts the number of service categories for which the physician billed. The five service categories included in the measure are: immunization, behavioral or mental health counseling, treatment of minor lacerations, cryotherapy/skin excision, and joint injection.

[i] To be included in the analysis of both long-term opioid use and potential overuse, a beneficiary had to: (1) be assigned to a practice; (2) be continuously enrolled in Medicare Parts A, B, and D throughout each calendar year or until death; and (3) have at least one opioid prescription during the measurement year. We further excluded beneficiaries for whom opioid use is appropriate: beneficiaries with a diagnosis of cancer during the measurement year or one year before, or a diagnosis of sickle cell disease or hospice use during the measurement year. The regression models for both opioid use outcomes additionally control for changes in state-level PDMP characteristics and opioid funding, and county-level opioid marketing intensity.

[j] This measure is defined only among long-term users of opioids.

[k] This measure is the annualized total number of services that provide little to no benefit to patients, have potential to cause harm, incur unnecessary costs to patients, or waste limited healthcare resources, per 1,000 beneficiaries. Because three of the low-value services are identified using a one-year look-forward period to determine whether the service was low-value or not, this measure is not available for PY 5.

*/**/*** Significantly different from zero at the 0.10/0.05/0.01 level, two-tailed test.

C = comparison; E&M = Evaluation and Management; FFS = fee-for-service; ITT = Intent-to-treat; NA = not applicable; NPI = National Provider Identifier; PDMP = prescription drug monitoring program; PY = Program Year; SE = standard error; SSP = Medicare Shared Savings Program.

Table 5.A.3.2b. Regression-adjusted means and estimated impacts of CPC+ on selected claims-based quality-of-care measures for attributed Medicare FFS beneficiaries by program year and average across the five program years, Track 2 by SSP status

	Track 2—SSP					Track 2—Non-SSP					
	CPC+ mean[a]	C mean[a]	Impact estimate[b] (SE)	90% confidence interval	*p*-Value	CPC+ mean[a]	C mean[a]	Impact estimate[b] (SE)	90% confidence interval	*p*-Value	*p*-Value for SSP vs. non-SSP difference
Planned care and population health measures for beneficiaries ages 18–75 with diabetes (percentage)											
Received HbA1c test											
Baseline	92.8%	92.0%	NA	NA	NA	92.2%	92.2%	NA	NA	NA	NA
PY 1	93.3%	92.4%	0.1 (0.2)	(-0.2, 0.5)	0.61	92.6%	92.2%	0.4 (0.3)	(0.0, 0.8)	0.13	0.41
PY 2	92.9%	92.1%	0.0 (0.3)	(-0.4, 0.5)	0.94	92.5%	92.1%	0.4 (0.3)	(-0.1, 0.8)	0.16	0.34
PY 3	92.7%	91.9%	0.0 (0.4)	(-0.6, 0.7)	0.94	92.7%	92.1%	0.6** (0.3)	(0.2, 1.1)	0.02	0.22
PY 4	89.9%	89.1%	0.0 (0.4)	(-0.6, 0.6)	0.95	89.9%	89.4%	0.5 (0.3)	(0.0, 1.0)	0.13	0.33
PY 5	92.3%	91.4%	0.1 (0.5)	(-0.7, 0.8)	0.89	92.1%	91.6%	0.5* (0.3)	(0.0, 1.0)	0.07	0.41
PY 1 through 5	92.2%	91.4%	0.0 (0.3)	(-0.4, 0.5)	0.88	91.9%	91.5%	0.5* (0.2)	(0.1, 0.9)	0.05	0.24
Received eye exam											
Baseline	66.9%	66.9%	NA	NA	NA	64.2%	64.4%	NA	NA	NA	NA
PY 1	66.9%	67.6%	-0.7* (0.4)	(-1.4, 0.0)	0.10	65.4%	65.3%	0.4 (0.3)	(0.0, 0.9)	0.13	0.03
PY 2	69.0%	67.6%	1.4*** (0.5)	(0.6, 2.2)	0.01	66.0%	65.2%	1.1*** (0.4)	(0.5, 1.7)	0.00	0.61
PY 3	70.1%	67.6%	2.6*** (0.6)	(1.6, 3.5)	0.00	66.3%	65.9%	0.6 (0.4)	(-0.1, 1.3)	0.13	0.01
PY 4	65.3%	63.1%	2.2*** (0.6)	(1.2, 3.3)	0.00	62.0%	61.8%	0.5 (0.4)	(-0.2, 1.2)	0.26	0.03
PY 5	67.6%	65.8%	1.8*** (0.7)	(0.7, 3.0)	0.01	64.5%	63.7%	1.0** (0.5)	(0.2, 1.8)	0.03	0.29
PY 1 through 5	67.8%	66.3%	1.5*** (0.5)	(0.7, 2.3)	0.00	64.8%	64.4%	0.7** (0.3)	(0.2, 1.3)	0.03	0.20
Received attention for nephropathy											
Baseline	84.6%	82.8%	NA	NA	NA	81.2%	81.6%	NA	NA	NA	NA
PY 1	85.2%	83.4%	0.0 (0.3)	(-0.6, 0.5)	0.97	81.9%	81.9%	0.5 (0.3)	(-0.1, 1.0)	0.14	0.29
PY 2	85.7%	83.7%	0.2 (0.4)	(-0.5, 0.8)	0.68	82.7%	81.9%	1.2*** (0.4)	(0.6, 1.9)	0.00	0.06
PY 3	85.0%	83.7%	-0.6 (0.4)	(-1.3, 0.1)	0.16	82.9%	82.2%	1.1** (0.4)	(0.4, 1.8)	0.01	0.01
PY 4	81.4%	80.9%	-1.3** (0.5)	(-2.1, -0.4)	0.02	78.9%	79.5%	-0.2 (0.5)	(-1.0, 0.6)	0.71	0.14

Table 5.A.3.2b. ***(continued)***

	Track 2—SSP					Track 2—Non-SSP					
	CPC+ mean[a]	C mean[a]	Impact estimate[b] (SE)	90% confidence interval	*p*-Value	CPC+ mean[a]	C mean[a]	Impact estimate[b] (SE)	90% confidence interval	*p*-Value	*p*-Value for SSP vs. non-SSP difference
PY 5	84.2%	82.8%	-0.4 (0.5)	(-1.3, 0.4)	0.38	81.2%	81.1%	0.5 (0.5)	(-0.2, 1.3)	0.24	0.14
PY 1 through 5	84.3%	82.8%	-0.4 (0.4)	(-1.0, 0.2)	0.27	81.5%	81.3%	0.7* (0.4)	(0.1, 1.3)	0.06	0.04
Diabetes Composite Measure 1 (received all three tests above: HbA1c test, eye exam, attention for nephropathy)											
Baseline	56.4%	55.2%	NA	NA	NA	52.1%	52.4%	NA	NA	NA	NA
PY 1	57.0%	56.2%	-0.5 (0.5)	(-1.2, 0.3)	0.28	53.6%	53.3%	0.6* (0.4)	(0.0, 1.2)	0.08	0.05
PY 2	59.1%	56.1%	1.8*** (0.5)	(0.9, 2.7)	0.00	54.4%	53.3%	1.4*** (0.4)	(0.7, 2.2)	0.00	0.65
PY 3	59.4%	56.2%	2.0*** (0.6)	(1.0, 3.0)	0.00	54.7%	54.2%	0.9* (0.5)	(0.1, 1.6)	0.08	0.13
PY 4	52.9%	50.4%	1.3* (0.7)	(0.2, 2.4)	0.05	49.0%	49.0%	0.3 (0.5)	(-0.6, 1.2)	0.55	0.26
PY 5	56.6%	53.8%	1.6** (0.7)	(0.4, 2.8)	0.03	52.2%	51.5%	1.0* (0.5)	(0.1, 1.9)	0.07	0.51
PY 1 through 5	57.0%	54.5%	1.2** (0.5)	(0.4, 2.1)	0.01	52.7%	52.2%	0.9** (0.4)	(0.2, 1.5)	0.03	0.55
Diabetes Composite Measure 2 (received none of the three tests above)											
Baseline	2.0%	2.2%	NA	NA	NA	2.2%	2.1%	NA	NA	NA	NA
PY 1	1.9%	2.0%	0.1 (0.1)	(-0.1, 0.3)	0.58	2.0%	2.1%	-0.2* (0.1)	(-0.4, 0.0)	0.09	0.11
PY 2	1.9%	2.0%	0.0 (0.1)	(-0.2, 0.2)	0.71	2.0%	2.2%	-0.3*** (0.1)	(-0.5, -0.1)	0.01	0.03
PY 3	1.9%	2.1%	0.0 (0.1)	(-0.2, 0.1)	0.77	1.9%	2.2%	-0.4*** (0.1)	(-0.6, -0.2)	0.00	0.03
PY 4	3.2%	3.1%	0.3** (0.1)	(0.0, 0.5)	0.05	3.3%	3.3%	-0.1 (0.1)	(-0.3, 0.2)	0.66	0.09
PY 5	2.1%	2.1%	0.2 (0.1)	(0.0, 0.4)	0.17	2.2%	2.4%	-0.3** (0.1)	(-0.5, -0.1)	0.01	0.01
PY 1 through 5	2.2%	2.3%	0.1 (0.1)	(-0.1, 0.3)	0.31	2.3%	2.5%	-0.3*** (0.1)	(-0.4, -0.1)	0.01	0.01
Unweighted sample sizes for the diabetes measures[c]											
Number of beneficiaries	142,668	378,418				183,560	386,413				
Number of beneficiary-years	419,013	1,111,621				545,348	1,131,904				

Table 5.A.3.2b. ***(continued)***

	Track 2—SSP					Track 2—Non-SSP					
	CPC+ mean[a]	C mean[a]	Impact estimate[b] (SE)	90% confidence interval	*p*-Value	CPC+ mean[a]	C mean[a]	Impact estimate[b] (SE)	90% confidence interval	*p*-Value	*p*-Value for SSP vs. non-SSP difference
Planned care and population health measures for female beneficiaries ages 52–74 (percentage)											
Received breast cancer screening											
Baseline	75.6%	75.0%	NA	NA	NA	72.0%	73.7%	NA	NA	NA	NA
PY 1	76.5%	75.7%	0.2 (0.2)	(-0.1, 0.6)	0.30	73.2%	74.2%	0.6*** (0.2)	(0.3, 1.0)	0.00	0.18
PY 2	77.2%	76.3%	0.3 (0.3)	(-0.1, 0.8)	0.24	74.0%	74.4%	1.3*** (0.3)	(0.8, 1.7)	0.00	0.02
PY 3	77.5%	76.7%	0.2 (0.3)	(-0.4, 0.7)	0.59	74.7%	74.9%	1.5*** (0.3)	(1.1, 2.0)	0.00	0.00
PY 4	75.0%	74.5%	0.0 (0.4)	(-0.7, 0.6)	0.93	73.0%	73.2%	1.4*** (0.3)	(0.9, 2.0)	0.00	0.00
PY 5	75.1%	74.8%	-0.3 (0.4)	(-0.9, 0.3)	0.39	73.1%	73.4%	1.4*** (0.4)	(0.8, 1.9)	0.00	0.00
PY 1 through 5	76.2%	75.5%	0.1 (0.3)	(-0.4, 0.6)	0.76	73.6%	74.0%	1.3*** (0.2)	(0.9, 1.7)	0.00	0.00
Unweighted sample sizes for the breast cancer screening measure[c]											
Number of beneficiaries	238,983	622,298				296,849	625,769				
Number of beneficiary-years	731,475	1,901,128				916,879	1,915,674				
Planned care and population health measures for beneficiaries ages 21 and older[d]											
Percentage of beneficiaries with cardiovascular disease who were prescribed and filled statin therapy											
Baseline	60.0%	59.7%	NA	NA	NA	58.9%	59.6%	NA	NA	NA	NA
PY 1	61.3%	61.2%	-0.2 (0.1)	(-0.4, 0.0)	0.18	60.1%	60.7%	0.2 (0.1)	(-0.1, 0.4)	0.21	0.07
PY 2	60.3%	60.2%	-0.3 (0.2)	(-0.6, 0.0)	0.11	59.3%	59.9%	0.1 (0.2)	(-0.2, 0.5)	0.49	0.11
PY 3	61.3%	61.6%	-0.7*** (0.2)	(-1.0, -0.3)	0.00	60.6%	61.0%	0.4 (0.2)	(0.0, 0.7)	0.11	0.00
PY 4	61.9%	62.4%	-0.8*** (0.2)	(-1.2, -0.4)	0.00	61.2%	61.7%	0.3 (0.2)	(-0.1, 0.7)	0.26	0.00
PY 5	63.3%	63.9%	-0.9*** (0.3)	(-1.4, -0.5)	0.00	62.6%	63.0%	0.4 (0.3)	(0.0, 0.8)	0.15	0.00
PY 1 through 5	61.6%	61.9%	-0.6*** (0.2)	(-0.9, -0.3)	0.00	60.8%	61.3%	0.3 (0.2)	(-0.1, 0.6)	0.17	0.00
Unweighted sample sizes for the statin therapy measure[c]											
Number of beneficiaries	453,563	1,204,376				539,332	1,156,990				
Number of beneficiary-years	1,516,078	4,035,595				1,812,238	3,838,823				

Table 5.A.3.2b. ***(continued)***

	Track 2—SSP					Track 2—Non-SSP					
	CPC+ mean[a]	C mean[a]	Impact estimate[b] (SE)	90% confidence interval	*p*-Value	CPC+ mean[a]	C mean[a]	Impact estimate[b] (SE)	90% confidence interval	*p*-Value	*p*-Value for SSP vs. non-SSP difference
Planned care and population health measures for beneficiaries ages 18 and older[d]											
Percentage of beneficiaries on diabetes medications with proportion of days covered by medication > 80%											
Baseline	78.3%	78.1%	NA	NA	NA	76.7%	77.1%	NA	NA	NA	NA
PY 1	79.5%	78.9%	0.3 (0.4)	(-0.3, 1.0)	0.38	77.8%	78.5%	-0.3 (0.3)	(-0.8, 0.3)	0.44	0.24
PY 2	80.1%	79.6%	0.3 (0.4)	(-0.4, 1.0)	0.45	78.9%	79.3%	0.0 (0.3)	(-0.5, 0.6)	0.98	0.57
PY 3	79.5%	79.3%	0.0 (0.4)	(-0.7, 0.7)	0.97	78.1%	78.9%	-0.4 (0.3)	(-0.9, 0.2)	0.27	0.46
PY 4	82.9%	82.2%	0.5 (0.4)	(-0.2, 1.1)	0.25	81.6%	81.8%	0.2 (0.3)	(-0.3, 0.8)	0.52	0.65
PY 5	81.6%	81.7%	-0.4 (0.4)	(-1.0, 0.3)	0.33	80.6%	81.1%	-0.1 (0.4)	(-0.7, 0.5)	0.79	0.59
PY 1 through 5	80.8%	80.4%	0.2 (0.3)	(-0.4, 0.7)	0.65	79.5%	80.0%	-0.1 (0.3)	(-0.6, 0.3)	0.69	0.55
Percentage of beneficiaries on renin-angiotensin system antagonists with proportion of days covered by medication > 80%											
Baseline	79.1%	79.2%	NA	NA	NA	78.3%	78.5%	NA	NA	NA	NA
PY 1	81.9%	81.7%	0.2 (0.2)	(-0.1, 0.6)	0.34	80.5%	80.7%	0.0 (0.2)	(-0.3, 0.4)	0.93	0.53
PY 2	82.5%	82.5%	0.0 (0.2)	(-0.3, 0.4)	0.90	81.5%	81.7%	0.1 (0.2)	(-0.3, 0.4)	0.78	0.92
PY 3	82.5%	83.0%	-0.5** (0.2)	(-0.8, -0.1)	0.04	81.5%	81.9%	-0.3 (0.2)	(-0.6, 0.1)	0.22	0.51
PY 4	84.4%	85.0%	-0.5** (0.2)	(-0.8, -0.1)	0.03	83.7%	84.1%	-0.2 (0.2)	(-0.5, 0.2)	0.47	0.30
PY 5	84.1%	84.9%	-0.7*** (0.2)	(-1.1, -0.4)	0.00	83.1%	84.0%	-0.7*** (0.2)	(-1.0, -0.3)	0.00	0.85
PY 1 through 5	83.2%	83.5%	-0.3 (0.2)	(-0.6, 0.0)	0.11	82.2%	82.6%	-0.2 (0.2)	(-0.5, 0.1)	0.25	0.71
Percentage of beneficiaries on statins with proportion of days covered by medication > 80%											
Baseline	76.8%	77.0%	NA	NA	NA	76.1%	76.1%	NA	NA	NA	NA
PY 1	77.0%	77.0%	0.2 (0.2)	(-0.1, 0.6)	0.28	76.2%	76.0%	0.2 (0.2)	(-0.1, 0.6)	0.25	0.99
PY 2	80.0%	80.4%	-0.2 (0.2)	(-0.5, 0.1)	0.36	79.7%	79.4%	0.3 (0.2)	(0.0, 0.6)	0.13	0.09
PY 3	80.1%	80.7%	-0.5** (0.2)	(-0.8, -0.1)	0.03	79.6%	79.5%	0.1 (0.2)	(-0.2, 0.4)	0.65	0.05
PY 4	83.0%	83.5%	-0.3 (0.2)	(-0.7, 0.0)	0.11	82.6%	82.8%	-0.2 (0.2)	(-0.6, 0.1)	0.26	0.72
PY 5	82.4%	83.1%	-0.5** (0.2)	(-0.9, -0.2)	0.02	81.9%	82.2%	-0.3* (0.2)	(-0.7, 0.0)	0.09	0.54

Table 5.A.3.2b. ***(continued)***

	Track 2—SSP					Track 2—Non-SSP					
	CPC+ mean[a]	C mean[a]	Impact estimate[b] (SE)	90% confidence interval	*p*-Value	CPC+ mean[a]	C mean[a]	Impact estimate[b] (SE)	90% confidence interval	*p*-Value	*p*-Value for SSP vs. non-SSP difference
PY 1 through 5	80.7%	81.1%	-0.3 (0.2)	(-0.6, 0.0)	0.13	80.2%	80.2%	0.0 (0.2)	(-0.3, 0.3)	0.95	0.24
Percentage of beneficiaries with both coronary artery disease (CAD) and diabetes who were prescribed and filled angiotensin-converting enzyme (ACE) inhibitor or angiotensin receptor blocker (ARB) therapy											
Baseline	78.8%	78.1%	NA	NA	NA	78.3%	78.1%	NA	NA	NA	NA
PY 1	78.4%	77.9%	-0.1 (0.4)	(-0.7, 0.5)	0.82	78.2%	78.0%	0.1 (0.4)	(-0.5, 0.7)	0.77	0.72
PY 2	77.7%	77.1%	0.0 (0.4)	(-0.6, 0.7)	0.92	77.7%	77.2%	0.3 (0.4)	(-0.4, 1.0)	0.52	0.69
PY 3	77.8%	77.3%	-0.2 (0.4)	(-0.9, 0.5)	0.66	77.7%	77.1%	0.4 (0.4)	(-0.3, 1.1)	0.33	0.31
PY 4	76.5%	76.1%	-0.2 (0.5)	(-1.0, 0.6)	0.63	76.4%	76.1%	0.1 (0.5)	(-0.7, 0.8)	0.87	0.64
PY 5	76.9%	76.7%	-0.3 (0.5)	(-1.1, 0.5)	0.49	76.6%	76.2%	0.2 (0.5)	(-0.6, 1.0)	0.67	0.43
PY 1 through 5	77.5%	77.0%	-0.2 (0.4)	(-0.7, 0.4)	0.66	77.3%	76.9%	0.2 (0.4)	(-0.4, 0.8)	0.52	0.45
Unweighted sample sizes for percentage of beneficiaries on diabetes medications with proportion of days covered by medication> 80%											
Number of beneficiaries	101,375	268,049				123,656	264,638				
Number of beneficiary-years	304,902	802,780				371,111	785,850				
Unweighted sample sizes for percentage of beneficiaries on renin-angiotensin system antagonists with proportion of days covered by medication > 80%											
Number of beneficiaries	306,518	814,161				373,320	800,075				
Number of beneficiary-years	961,480	2,551,375				1,173,510	2,488,598				
Unweighted sample sizes for percentage of beneficiaries on statins with proportion of days covered by medication > 80%											
Number of beneficiaries	358,597	960,575				423,703	919,115				
Number of beneficiary-years	1,173,220	3,145,461				1,386,682	2,978,899				
Unweighted sample sizes for percentage of beneficiaries with both CAD and diabetes who were prescribed and filled ACE inhibitor or ARB therapy											
Number of beneficiaries	81,029	211,697				98,541	203,755				
Number of beneficiary-years	181,064	470,250				222,718	446,475				

Table 5.A.3.2b. ***(continued)***

	Track 2—SSP					Track 2—Non-SSP					
	CPC+ mean[a]	C mean[a]	Impact estimate[b] (SE)	90% confidence interval	*p*-Value	CPC+ mean[a]	C mean[a]	Impact estimate[b] (SE)	90% confidence interval	*p*-Value	*p*-Value for SSP vs. non-SSP difference
Measures for continuity of care[e]											
Percentage of primary care ambulatory visits at assigned practice											
Baseline	74.7%	73.9%	NA	NA	NA	76.0%	72.8%	NA	NA	NA	NA
PY 1	72.4%	71.1%	0.5 (0.4)	(-0.2, 1.1)	0.23	73.3%	69.8%	0.3 (0.3)	(-0.3, 0.8)	0.41	0.71
PY 2	64.0%	62.1%	1.1 (0.7)	(0.0, 2.2)	0.10	65.3%	61.2%	0.9* (0.5)	(0.0, 1.7)	0.09	0.81
PY 3	62.0%	59.4%	1.8** (0.7)	(0.6, 3.0)	0.01	63.2%	58.5%	1.5** (0.6)	(0.5, 2.5)	0.01	0.73
PY 4	54.4%	53.4%	0.1 (1.1)	(-1.8, 2.0)	0.94	56.8%	52.0%	1.6 (1.0)	(0.0, 3.3)	0.11	0.30
PY 5	49.9%	48.8%	0.2 (1.1)	(-1.7, 2.1)	0.89	52.5%	48.3%	1.0 (0.9)	(-0.5, 2.6)	0.28	0.57
PY 1 through 5	60.1%	58.4%	0.8 (0.7)	(-0.3, 1.9)	0.26	61.8%	57.5%	1.1** (0.5)	(0.2, 1.9)	0.05	0.73
Across all PCPs and specialists providing care to a patient, where each practitioner in the beneficiary's assigned practice is treated separately											
Percentage of visits with the usual provider of care (UPC)											
Baseline	46.9%	47.3%	NA	NA	NA	48.9%	48.4%	NA	NA	NA	NA
PY 1	46.0%	46.4%	0.0 (0.1)	(-0.2, 0.1)	0.82	47.8%	47.4%	-0.1 (0.1)	(-0.2, 0.1)	0.58	0.82
PY 2	45.0%	45.5%	-0.1 (0.1)	(-0.3, 0.1)	0.37	46.8%	46.4%	-0.1 (0.1)	(-0.3, 0.1)	0.29	0.93
PY 3	44.2%	44.7%	-0.1 (0.1)	(-0.3, 0.1)	0.61	46.0%	45.7%	-0.2 (0.1)	(-0.4, 0.0)	0.15	0.54
PY 4	46.8%	47.4%	-0.2 (0.2)	(-0.5, 0.1)	0.28	48.4%	48.2%	-0.3** (0.1)	(-0.6, -0.1)	0.04	0.59
PY 5	43.9%	44.5%	-0.2 (0.2)	(-0.5, 0.0)	0.16	45.7%	45.5%	-0.3** (0.1)	(-0.6, -0.1)	0.03	0.68
PY 1 through 5	45.1%	45.7%	-0.1 (0.1)	(-0.3, 0.1)	0.30	46.9%	46.6%	-0.2* (0.1)	(-0.4, 0.0)	0.07	0.64
Reversed Bice-Boxerman fragmentation of care index											
Baseline	78.7%	78.3%	NA	NA	NA	76.8%	77.6%	NA	NA	NA	NA
PY 1	79.7%	79.2%	0.1 (0.1)	(-0.1, 0.2)	0.59	77.8%	78.5%	0.1 (0.1)	(-0.1, 0.2)	0.38	0.81
PY 2	80.7%	80.1%	0.2 (0.1)	(0.0, 0.4)	0.15	78.8%	79.5%	0.1 (0.1)	(-0.1, 0.3)	0.32	0.74
PY 3	81.5%	81.0%	0.1 (0.1)	(-0.1, 0.3)	0.45	79.6%	80.2%	0.2 (0.1)	(-0.1, 0.4)	0.23	0.78
PY 4	81.6%	81.2%	0.0 (0.2)	(-0.3, 0.3)	0.97	79.9%	80.4%	0.2 (0.2)	(0.0, 0.5)	0.16	0.36

Table 5.A.3.2b. ***(continued)***

	Track 2—SSP					Track 2—Non-SSP					
	CPC+ mean[a]	C mean[a]	Impact estimate[b] (SE)	90% confidence interval	*p*-Value	CPC+ mean[a]	C mean[a]	Impact estimate[b] (SE)	90% confidence interval	*p*-Value	*p*-Value for SSP vs. non-SSP difference
PY 5	82.7%	82.1%	0.2 (0.2)	(-0.1, 0.5)	0.24	81.1%	81.5%	0.4** (0.2)	(0.1, 0.7)	0.01	0.38
PY 1 through 5	81.3%	80.8%	0.1 (0.1)	(-0.1, 0.3)	0.37	79.5%	80.1%	0.2* (0.1)	(0.0, 0.4)	0.09	0.61
Across all PCPs and specialists providing care to a patient, where all practitioners in the beneficiary's assigned practice are treated as a single practitioner											
Percentage of visits with the usual provider of care (UPC)											
Baseline	50.1%	50.6%	NA	NA	NA	52.3%	51.4%	NA	NA	NA	NA
PY 1	49.2%	49.6%	0.1 (0.1)	(-0.1, 0.3)	0.30	51.1%	50.3%	0.0 (0.1)	(-0.2, 0.2)	0.91	0.49
PY 2	47.6%	48.1%	0.0 (0.1)	(-0.3, 0.2)	0.87	49.4%	48.7%	-0.1 (0.1)	(-0.3, 0.1)	0.58	0.79
PY 3	47.7%	47.9%	0.2 (0.2)	(-0.1, 0.6)	0.24	49.3%	48.6%	-0.1 (0.1)	(-0.4, 0.1)	0.34	0.13
PY 4	49.4%	50.0%	-0.1 (0.2)	(-0.5, 0.3)	0.75	51.0%	50.5%	-0.3* (0.2)	(-0.6, 0.0)	0.09	0.42
PY 5	46.5%	47.0%	0.1 (0.2)	(-0.3, 0.5)	0.77	48.1%	47.7%	-0.4** (0.2)	(-0.7, -0.1)	0.03	0.12
PY 1 through 5	48.0%	48.5%	0.1 (0.2)	(-0.2, 0.3)	0.67	49.7%	49.1%	-0.2 (0.1)	(-0.4, 0.0)	0.14	0.22
Reversed Bice-Boxerman fragmentation of care index											
Baseline	75.4%	74.7%	NA	NA	NA	73.0%	74.2%	NA	NA	NA	NA
PY 1	76.3%	75.8%	-0.2 (0.1)	(-0.4, 0.1)	0.25	74.2%	75.4%	0.0 (0.1)	(-0.2, 0.2)	0.87	0.32
PY 2	78.0%	77.3%	0.0 (0.2)	(-0.2, 0.3)	0.89	76.0%	77.1%	0.1 (0.1)	(-0.1, 0.3)	0.46	0.69
PY 3	77.9%	77.5%	-0.3 (0.2)	(-0.7, 0.1)	0.24	76.1%	77.1%	0.1 (0.2)	(-0.1, 0.4)	0.41	0.15
PY 4	78.9%	78.3%	-0.1 (0.3)	(-0.6, 0.3)	0.66	77.1%	77.9%	0.3 (0.2)	(-0.1, 0.7)	0.18	0.23
PY 5	80.1%	79.5%	-0.1 (0.3)	(-0.6, 0.3)	0.58	78.6%	79.2%	0.6** (0.2)	(0.2, 0.9)	0.01	0.04
PY 1 through 5	78.3%	77.8%	-0.1 (0.2)	(-0.5, 0.2)	0.47	76.5%	77.4%	0.2 (0.1)	(0.0, 0.5)	0.12	0.13
Unweighted sample sizes for percentage of primary care ambulatory visits at assigned practice[c]											
Number of beneficiaries	722,520	1,932,581				899,528	1,929,482				
Number of beneficiary-years	2,532,383	6,771,294				3,185,249	6,707,885				

Table 5.A.3.2b. ***(continued)***

	Track 2—SSP					Track 2—Non-SSP					
	CPC+ mean[a]	C mean[a]	Impact estimate[b] (SE)	90% confidence interval	*p*-Value	CPC+ mean[a]	C mean[a]	Impact estimate[b] (SE)	90% confidence interval	*p*-Value	*p*-Value for SSP vs. non-SSP difference
Unweighted sample sizes for percentage of visits with the usual provider of care[c]											
Number of beneficiaries	736,713	1,974,760				916,565	1,971,742				
Number of beneficiary-years	2,656,325	7,120,953				3,331,828	7,061,859				
Unweighted sample sizes for reversed Bice-Boxerman fragmentation of care index[c]											
Number of beneficiaries	667,640	1,797,923				829,283	1,779,670				
Number of beneficiary-years	2,168,235	5,841,539				2,711,476	5,707,536				
Comprehensiveness of care (measured at the physician level)											
Physician involvement in patient conditions[f]											
Baseline	66.0%	66.5%	NA	NA	NA	68.0%	68.0%	NA	NA	NA	NA
PY 1	65.9%	66.7%	-0.3 (0.4)	(-1.0, 0.4)	0.45	67.8%	67.7%	0.0 (0.4)	(-0.6, 0.7)	0.93	0.57
PY 2	65.6%	66.7%	-0.6 (0.4)	(-1.4, 0.1)	0.15	68.0%	67.6%	0.4 (0.4)	(-0.3, 1.1)	0.40	0.11
PY 3	66.5%	67.5%	-0.5 (0.5)	(-1.4, 0.3)	0.31	68.9%	68.6%	0.3 (0.5)	(-0.5, 1.1)	0.57	0.27
PY 4	69.2%	70.6%	-1.0 (0.6)	(-2.0, 0.0)	0.11	71.4%	70.4%	1.0* (0.6)	(0.0, 1.9)	0.10	0.02
PY 5	66.8%	68.7%	-1.4** (0.6)	(-2.4, -0.4)	0.02	69.3%	68.2%	1.0* (0.6)	(0.0, 2.0)	0.10	0.01
PY 1 through 5	66.7%	67.9%	-0.7* (0.4)	(-1.3, -0.1)	0.07	69.0%	68.5%	0.4 (0.3)	(-0.1, 1.0)	0.20	0.03
Range of services provided by physicians[h]											
Baseline	1.91	1.90	NA	NA	NA	2.14	1.93	NA	NA	NA	NA
PY 1	1.85	1.85	-0.01 (0.02)	(-0.05, 0.03)	0.67	2.11	1.90	0.01 (0.03)	(-0.03, 0.05)	0.78	0.62
PY 2	1.81	1.80	0.00 (0.03)	(-0.04, 0.05)	0.89	2.10	1.85	0.05* (0.03)	(0.00, 0.10)	0.09	0.26
PY 3	1.82	1.77	0.04 (0.03)	(-0.02, 0.09)	0.28	2.11	1.86	0.05 (0.03)	(-0.01, 0.10)	0.16	0.83
PY 4	1.71	1.74	-0.04 (0.04)	(-0.10, 0.03)	0.34	2.02	1.81	0.01 (0.04)	(-0.06, 0.07)	0.88	0.44
PY 5	1.68	1.69	-0.02 (0.04)	(-0.08, 0.04)	0.62	1.97	1.73	0.03 (0.04)	(-0.04, 0.09)	0.47	0.39
PY 1 through 5	1.78	1.77	0.00 (0.02)	(-0.04, 0.04)	0.92	2.07	1.84	0.03 (0.03)	(-0.01, 0.07)	0.25	0.37

Table 5.A.3.2b. ***(continued)***

	Track 2—SSP					Track 2—Non-SSP					
	CPC+ mean[a]	C mean[a]	Impact estimate[b] (SE)	90% confidence interval	*p*-Value	CPC+ mean[a]	C mean[a]	Impact estimate[b] (SE)	90% confidence interval	*p*-Value	*p*-Value for SSP vs. non-SSP difference
Management of new problems by physicians[g]											
Baseline	0.994	0.995	NA	NA	NA	1.000	0.998	NA	NA	NA	NA
PY 1	0.991	0.995	-0.003 (0.003)	(-0.008, 0.003)	0.40	0.998	1.000	-0.004 (0.003)	(-0.009, 0.000)	0.14	0.73
PY 2	0.991	0.996	-0.004 (0.003)	(-0.009, 0.002)	0.30	0.998	0.997	-0.001 (0.003)	(-0.007, 0.004)	0.66	0.64
PY 3	0.993	0.993	0.002 (0.003)	(-0.004, 0.007)	0.59	0.997	1.001	-0.006* (0.003)	(-0.012, -0.001)	0.06	0.09
PY 4	0.992	0.995	-0.002 (0.004)	(-0.008, 0.005)	0.66	0.994	0.998	-0.006* (0.004)	(-0.013, 0.000)	0.08	0.40
PY 1 through 4	0.992	0.995	-0.002 (0.003)	(-0.006, 0.003)	0.53	0.997	0.999	-0.004* (0.003)	(-0.008, 0.000)	0.09	0.50
Unweighted sample sizes for Physician involvement in patient conditions											
Number of physicians	2,561	6,708				3,238	7,193				
Number of physician-years	12,184	30,897				15,466	32,642				
Unweighted sample sizes for Range of services provided by physicians[g]											
Number of physicians	2,448	6,439				3,135	6,836				
Number of physician-years	11,755	29,370				14,984	30,645				
Unweighted sample sizes for Management of new problems by physicians											
Number of physicians	2,488	6,534				3,165	7,020				
Number of physician-years	10,263	26,235				13,111	27,559				
Patient and caregiver engagement											
Percentage of beneficiaries receiving hospice services											
Baseline	2.7%	2.6%	NA	NA	NA	2.8%	2.8%	NA	NA	NA	NA
PY 1	2.7%	2.6%	0.0 (0.0)	(-0.1, 0.1)	0.77	2.8%	2.8%	0.1* (0.0)	(0.0, 0.1)	0.07	0.30
PY 2	2.9%	2.8%	0.1* (0.0)	(0.0, 0.2)	0.09	3.0%	2.9%	0.2*** (0.0)	(0.1, 0.3)	0.00	0.10
PY 3	3.1%	2.9%	0.1** (0.1)	(0.0, 0.2)	0.02	3.2%	3.0%	0.2*** (0.0)	(0.1, 0.3)	0.00	0.29
PY 4	3.2%	3.1%	0.1 (0.1)	(0.0, 0.1)	0.32	3.3%	3.2%	0.2*** (0.1)	(0.1, 0.2)	0.00	0.18

Table 5.A.3.2b. ***(continued)***

	Track 2—SSP					Track 2—Non-SSP					
	CPC+ mean[a]	C mean[a]	Impact estimate[b] (SE)	90% confidence interval	*p*-Value	CPC+ mean[a]	C mean[a]	Impact estimate[b] (SE)	90% confidence interval	*p*-Value	*p*-Value for SSP vs. non-SSP difference
PY 5	3.2%	3.1%	0.1* (0.1)	(0.0, 0.2)	0.07	3.3%	3.1%	0.2*** (0.1)	(0.1, 0.3)	0.00	0.15
PY 1 through 5	3.0%	2.9%	0.1* (0.0)	(0.0, 0.1)	0.07	3.2%	3.0%	0.2*** (0.0)	(0.1, 0.2)	0.00	0.11
Length of hospice stay, in days (for beneficiaries receiving hospice services)											
Baseline	59	64	NA	NA	NA	65	70	NA	NA	NA	NA
PY 1	59	65	0.3 (1.6)	(-2.3, 2.9)	0.86	64	67	0.8 (1.3)	(-1.4, 3.0)	0.56	0.81
PY 2	63	66	2.8* (1.7)	(0.0, 5.5)	0.10	68	71	2.2 (1.5)	(-0.3, 4.7)	0.15	0.80
PY 3	69	70	4.8*** (1.6)	(2.2, 7.5)	0.00	73	74	3.4** (1.6)	(0.7, 6.0)	0.04	0.52
PY 4	68	70	3.6** (1.7)	(0.8, 6.5)	0.03	70	73	1.5 (1.7)	(-1.3, 4.2)	0.37	0.36
PY 5	69	71	3.0* (1.7)	(0.1, 5.8)	0.08	71	74	1.4 (1.6)	(-1.3, 4.0)	0.39	0.50
PY 1 through 5	66	69	3.0** (1.4)	(0.8, 5.3)	0.03	70	72	1.9 (1.3)	(-0.2, 4.0)	0.14	0.56
Length of hospice stay, in days (for all beneficiaries)											
Baseline	1.6	1.7	NA	NA	NA	1.8	2.0	NA	NA	NA	NA
PY 1	1.6	1.7	0.0 (0.0)	(-0.1, 0.1)	0.67	1.8	1.9	0.1 (0.0)	(0.0, 0.1)	0.12	0.42
PY 2	1.8	1.8	0.1** (0.1)	(0.0, 0.2)	0.02	2.1	2.0	0.2*** (0.1)	(0.1, 0.3)	0.00	0.41
PY 3	2.1	2.0	0.2*** (0.1)	(0.1, 0.3)	0.00	2.3	2.2	0.2*** (0.1)	(0.1, 0.3)	0.00	0.95
PY 4	2.2	2.1	0.2*** (0.1)	(0.1, 0.3)	0.01	2.4	2.3	0.2** (0.1)	(0.0, 0.3)	0.02	0.88
PY 5	2.2	2.1	0.2*** (0.1)	(0.1, 0.3)	0.01	2.4	2.3	0.2*** (0.1)	(0.1, 0.3)	0.00	0.85
PY 1 through 5	2.0	2.0	0.2*** (0.1)	(0.1, 0.2)	0.00	2.2	2.2	0.2*** (0.1)	(0.1, 0.3)	0.00	0.74
Unweighted sample sizes for patient and caregiver engagement measures											
Number of beneficiaries for length of hospice stay	76,398	194,122				98,527	196,364				

Table 5.A.3.2b. ***(continued)***

	Track 2—SSP					Track 2—Non-SSP					
	CPC+ mean[a]	C mean[a]	Impact estimate[b] (SE)	90% confidence interval	*p*-Value	CPC+ mean[a]	C mean[a]	Impact estimate[b] (SE)	90% confidence interval	*p*-Value	*p*-Value for SSP vs. non-SSP difference
Other quality measures											
Percentage of index acute care hospital discharges that were followed by an unplanned readmission within 30 days											
Baseline	15.8%	15.9%	NA	NA	NA	15.5%	15.7%	NA	NA	NA	NA
PY 1	16.0%	16.1%	0.0 (0.2)	(-0.4, 0.3)	0.89	15.3%	15.6%	-0.1 (0.2)	(-0.5, 0.3)	0.67	0.84
PY 2	16.2%	16.0%	0.3 (0.2)	(-0.1, 0.6)	0.18	15.5%	16.0%	-0.2 (0.2)	(-0.6, 0.2)	0.42	0.13
PY 3	16.2%	16.2%	0.1 (0.2)	(-0.3, 0.5)	0.78	15.7%	16.2%	-0.2 (0.2)	(-0.6, 0.1)	0.30	0.37
PY 4	16.4%	16.0%	0.5** (0.2)	(0.2, 0.9)	0.02	15.9%	16.2%	-0.1 (0.2)	(-0.5, 0.3)	0.74	0.06
PY 5	16.6%	16.1%	0.6** (0.2)	(0.2, 1.0)	0.01	16.0%	16.1%	0.1 (0.2)	(-0.3, 0.5)	0.58	0.17
PY 1 through 5	16.3%	16.1%	0.3 (0.2)	(0.0, 0.6)	0.13	15.7%	16.0%	-0.1 (0.2)	(-0.4, 0.2)	0.57	0.15
Percentage of index acute care hospital discharges that were followed by an unplanned acute care hospitalization or ED visit (including observation stays) within 30 days											
Baseline	25.7%	25.9%	NA	NA	NA	25.9%	26.2%	NA	NA	NA	NA
PY 1	25.9%	26.1%	-0.1 (0.3)	(-0.5, 0.3)	0.75	25.9%	26.2%	0.0 (0.2)	(-0.4, 0.4)	0.91	0.88
PY 2	26.0%	26.1%	0.2 (0.3)	(-0.3, 0.6)	0.54	26.1%	26.5%	-0.2 (0.2)	(-0.6, 0.3)	0.54	0.39
PY 3	26.0%	26.2%	0.0 (0.3)	(-0.4, 0.4)	0.98	26.4%	26.7%	-0.1 (0.2)	(-0.5, 0.3)	0.60	0.73
PY 4	25.4%	25.5%	0.1 (0.3)	(-0.3, 0.6)	0.57	25.8%	26.1%	0.0 (0.3)	(-0.5, 0.4)	0.89	0.61
PY 5	25.7%	25.8%	0.1 (0.3)	(-0.3, 0.6)	0.66	26.0%	26.2%	0.0 (0.3)	(-0.4, 0.5)	0.85	0.85
PY 1 through 5	25.8%	26.0%	0.1 (0.2)	(-0.3, 0.4)	0.75	26.0%	26.4%	-0.1 (0.2)	(-0.4, 0.3)	0.73	0.64
Percentage of index ED (including observation stays) discharges that were followed by an unplanned acute care hospitalization or ED visit (including observation stays) within 30 days											
Baseline	28.6%	28.9%	NA	NA	NA	29.9%	30.4%	NA	NA	NA	NA
PY 1	28.3%	28.8%	-0.1 (0.3)	(-0.6, 0.3)	0.60	29.6%	29.9%	0.1 (0.2)	(-0.3, 0.5)	0.61	0.47
PY 2	28.5%	28.5%	0.3 (0.3)	(-0.1, 0.8)	0.20	29.5%	30.3%	-0.3 (0.2)	(-0.7, 0.1)	0.25	0.08
PY 3	28.8%	28.5%	0.6** (0.3)	(0.1, 1.1)	0.04	29.6%	30.4%	-0.3 (0.3)	(-0.7, 0.2)	0.29	0.03
PY 4	29.2%	28.8%	0.7* (0.3)	(0.1, 1.2)	0.05	30.0%	30.2%	0.2 (0.3)	(-0.3, 0.7)	0.55	0.28
PY 5	28.6%	28.5%	0.4 (0.3)	(-0.1, 1.0)	0.19	29.3%	29.4%	0.4 (0.3)	(-0.1, 0.9)	0.17	0.97

Table 5.A.3.2b. ***(continued)***

	Track 2—SSP					Track 2—Non-SSP					
	CPC+ mean[a]	C mean[a]	Impact estimate[b] (SE)	90% confidence interval	*p*-Value	CPC+ mean[a]	C mean[a]	Impact estimate[b] (SE)	90% confidence interval	*p*-Value	*p*-Value for SSP vs. non-SSP difference
PY 1 through 5	28.7%	28.6%	0.4 (0.3)	(0.0, 0.8)	0.14	29.6%	30.1%	0.0 (0.2)	(-0.4, 0.3)	0.96	0.25
Percentage of 65 and older Medicare FFS beneficiaries who received two or more prescriptions for high risk medications in the same class[d]											
Baseline	11.6%	11.1%	NA	NA	NA	12.1%	12.3%	NA	NA	NA	NA
PY 1	11.9%	11.5%	-0.1 (0.1)	(-0.2, 0.1)	0.41	12.3%	12.6%	0.0 (0.1)	(-0.2, 0.1)	0.92	0.62
PY 2	11.7%	11.3%	-0.1 (0.1)	(-0.2, 0.1)	0.35	12.2%	12.4%	0.0 (0.1)	(-0.2, 0.1)	0.74	0.66
PY 3	13.9%	13.4%	0.1 (0.1)	(-0.1, 0.3)	0.60	14.2%	14.4%	0.0 (0.1)	(-0.2, 0.2)	0.79	0.85
PY 4	13.7%	13.2%	0.0 (0.1)	(-0.2, 0.2)	0.97	14.2%	14.3%	0.2 (0.1)	(0.0, 0.4)	0.18	0.36
PY 5	13.4%	12.8%	0.0 (0.1)	(-0.2, 0.3)	0.74	13.8%	14.0%	0.1 (0.1)	(-0.1, 0.3)	0.44	0.76
PY 1 through 5	13.0%	12.5%	0.0 (0.1)	(-0.2, 0.2)	0.95	13.4%	13.6%	0.1 (0.1)	(-0.1, 0.2)	0.59	0.67
Long-term opioid use[i]											
Baseline	8.0%	7.7%	NA	NA	NA	9.5%	9.0%	NA	NA	NA	NA
PY 1	7.3%	7.1%	-0.1 (0.1)	(-0.2, 0.1)	0.42	8.9%	8.2%	0.2** (0.1)	(0.0, 0.3)	0.03	0.03
PY 2	6.7%	6.5%	-0.1 (0.1)	(-0.3, 0.1)	0.43	8.1%	7.5%	0.2 (0.1)	(0.0, 0.3)	0.11	0.11
PY 3	5.9%	5.8%	-0.2 (0.1)	(-0.4, 0.0)	0.20	7.2%	6.7%	0.0 (0.1)	(-0.2, 0.2)	0.93	0.36
PY 4	5.4%	5.3%	-0.2 (0.1)	(-0.4, 0.1)	0.23	6.5%	6.1%	-0.1 (0.1)	(-0.2, 0.1)	0.65	0.51
PY 5	4.8%	4.8%	-0.2 (0.1)	(-0.5, 0.0)	0.12	5.8%	5.4%	-0.1 (0.1)	(-0.3, 0.1)	0.52	0.42
PY 1 through 5	6.0%	5.8%	-0.1 (0.1)	(-0.3, 0.0)	0.20	7.2%	6.7%	0.1 (0.1)	(-0.1, 0.2)	0.54	0.18
Potential opioid overuse[j]											
Baseline	19.4%	19.1%	NA	NA	NA	19.5%	19.4%	NA	NA	NA	NA
PY 1	18.2%	17.4%	0.5 (0.4)	(-0.1, 1.1)	0.18	17.3%	17.3%	-0.2 (0.3)	(-0.7, 0.3)	0.55	0.16
PY 2	16.4%	16.5%	-0.4 (0.5)	(-1.2, 0.4)	0.41	15.0%	15.2%	-0.4 (0.4)	(-1.1, 0.3)	0.31	0.96
PY 3	14.7%	14.6%	-0.2 (0.6)	(-1.1, 0.7)	0.68	12.7%	13.6%	-1.1** (0.5)	(-1.9, -0.3)	0.02	0.22
PY 4	13.5%	14.0%	-0.8 (0.6)	(-1.8, 0.2)	0.17	11.8%	12.4%	-0.7 (0.5)	(-1.5, 0.1)	0.15	0.86
PY 5	12.9%	13.2%	-0.7 (0.6)	(-1.7, 0.4)	0.29	11.0%	11.5%	-0.7 (0.5)	(-1.6, 0.2)	0.19	0.94

Table 5.A.3.2b. ***(continued)***

	Track 2—SSP					Track 2—Non-SSP					
	CPC+ mean[a]	C mean[a]	Impact estimate[b] (SE)	90% confidence interval	*p*-Value	CPC+ mean[a]	C mean[a]	Impact estimate[b] (SE)	90% confidence interval	*p*-Value	*p*-Value for SSP vs. non-SSP difference
PY 1 through 5	15.3%	15.1%	-0.1 (0.4)	(-0.8, 0.5)	0.73	13.7%	14.1%	-0.5 (0.3)	(-1.1, 0.0)	0.12	0.44
Annualized number of low-value services per 1,000 beneficiaries[k]											
Baseline	350	355	NA	NA	NA	340	339	NA	NA	NA	NA
PY 1	317	324	-2.4 (2.7)	(-6.9, 2.1)	0.37	311	309	1.1 (2.3)	(-2.7, 4.9)	0.64	0.33
PY 2	316	325	-4.2 (2.9)	(-9.1, 0.6)	0.15	314	313	0.8 (2.9)	(-4.0, 5.6)	0.79	0.23
PY 3	319	329	-6.0 (4.0)	(-12.5, 0.6)	0.13	320	316	3.3 (3.4)	(-2.4, 8.9)	0.34	0.08
PY 4	267	279	-7.2* (3.9)	(-13.7, -0.8)	0.06	272	267	4.7 (3.6)	(-1.2, 10.6)	0.19	0.02
PY 1 through 4	304	313	-5.0* (3.0)	(-9.9, -0.1)	0.09	303	301	2.4 (2.7)	(-2.0, 6.9)	0.37	0.07
Unweighted sample sizes for other quality of care measures											
Number of index discharges for readmission	779,836	1,998,608				916,771	1,958,908				
Number of index ED discharges	1,302,059	3,490,169				1,732,851	3,750,424				
Number of 65 and older Medicare FFS beneficiaries for the high-risk medication measure	542,136	1,424,555				640,766	1,377,798				
Number of beneficiaries for long-term opioid use	542,941	1,440,270				664,340	1,427,908				
Number of beneficiaries for potential opioid overuse	47,489	121,415				69,576	134,677				
Number of beneficiaries for low-value services measure	683,029	1,816,283				847,489	1,818,185				

Source: Mathematica's analysis of Medicare claims data from January 2013 through December 2021.

Notes: For the quality-of-care outcomes, we present the absolute impact estimate only. We do so because percentage impacts for some of the binary outcomes are likely to be misleadingly large, given the low means for the outcome measures.

This table indicates which estimates are statistically significant; when we interpret evidence, we combine evidence from the magnitude of the effect, the *p*-values, findings on related outcomes, subgroups, sensitivity tests, and other data sources on model implementation.

Table 5.A.3.2b. ***(continued)***

We grouped the claims-based quality-of-care measures into separate domains according to the Comprehensive Primary Care Functions under which they appear in the 2018 CPC+ Implementation Guide (CMMI 2018).

[a] We report the actual, unadjusted averages in the baseline period which are similar for the CPC+ and comparison groups due to matching. In the intervention periods, the comparison group mean is computed by subtracting the regression adjusted difference between the CPC+ and comparison means in each time period from the CPC+ mean in that same time period.

[b] Each impact estimate is regression-adjusted using a difference-in-differences analysis that reflects the difference of the average outcome for Medicare FFS beneficiaries attributed to CPC+ practices in the five years of CPC+ to the average outcome in the baseline year, relative to the same difference over time for Medicare FFS beneficiaries attributed to comparison practices, while controlling for beneficiary characteristics and practice fixed effects.

[c] The numbers of Track 2 CPC+ and comparison practices are same as in Table 5.A.2.1b and hence, are not reported separately in this table. The beneficiary-level measures for recommended services for diabetes, breast cancer screening, continuity of care, and opioid use are affected only by matching weights (and not by time observed) because the measures require beneficiaries to have full year of eligibility in each program year. After accounting for matching weights, the effective sample size for the comparison group for the measures presented in this table is 37 to 45 percent of the size of the actual comparison group.

[d] These measures require that beneficiaries be continuously enrolled in Medicare FFS Parts A and B as well as in Medicare Part D, and not use hospice services during the measurement year.

[e] The continuity of care measures are calculated for beneficiaries who were in the ITT sample at the beginning of the year and were FFS eligible for the full year in each program year and had qualifying ambulatory visits in the program year. Qualifying ambulatory visits are (1) office or other outpatient visit for E&M; (2) ophthalmological services; (3) medical examination and evaluation; and (4) new enrollee and annual wellness visits.

[f] For each physician, this measure indicates the percentage of beneficiaries for whom the physician was considered “most comprehensive” (i.e. saw the beneficiary for the largest share of their unique diagnoses codes) out of all beneficiaries the physician saw in the year.

[g] The new problem management measure is a score that indicates how often a primary care physician continues to treat a beneficiary’s new condition versus referring the beneficiary (or the beneficiary self-referring) to a specialist or different provider. Since the new problem management measure requires a one-year look forward period, this measure is not available for PY 5 (as creating the measure for PY 5 would have required using incomplete 2022 claims data).

[h] The range of services measure is a score ranging from 0–5 that counts the number of service categories for which the physician billed. The five service categories included in the measure are: immunization, behavioral or mental health counseling, treatment of minor lacerations, cryotherapy/skin excision, and joint injection.

[i] To be included in the analysis of both long-term opioid use and potential overuse, a beneficiary had to: (1) be assigned to a practice; (2) be continuously enrolled in Medicare Parts A, B, and D throughout each calendar year or until death; and (3) have at least one opioid prescription during the measurement year. We further excluded beneficiaries for whom opioid use is appropriate: beneficiaries with a diagnosis of cancer during the measurement year or one year before, or a diagnosis of sickle cell disease or hospice use during the measurement year. The regression models for both opioid use outcomes additionally control for changes in state-level PDMP characteristics and opioid funding, and county-level opioid marketing intensity.

[j] This measure is defined only among long-term users of opioids.

[k] This measure is the annualized total number of services that provide little to no benefit to patients, have potential to cause harm, incur unnecessary costs to patients, or waste limited healthcare resources, per 1,000 beneficiaries. Because three of the low-value services are identified using a one-year look-forward period to determine whether the service was low-value or not, this measure is not available for PY 5.

*/**/*** Significantly different from zero at the 0.10/0.05/0.01 level, two-tailed test.

C = comparison; E&M = Evaluation and Management; FFS = fee-for-service; ITT = Intent-to-treat; NA = not applicable; NPI = National Provider Identifier; PDMP = prescription drug monitoring program; PY = Program Year; SE = standard error; SSP = Medicare Shared Savings Program.

5.A.4. Aggregate impact estimates for key outcomes

Table 5.A.4.1. Aggregate impact estimates for key outcomes across the five years of CPC+, Track 1

Outcome	Estimate	90 percent CI lower bound	90 percent CI upper bound
Medicare expenditures without CMS's enhanced payments[a]	$63,309,345	-$244,444,548	$371,063,238
Medicare expenditures including CMS's enhanced payments[a]	***$721,187,847***	$415,930,250	$1,026,445,444
Hospitalizations	***-11,815***	-22,983	-646
Outpatient ED visits	***-44,409***	-63,707	-25,111
30-day readmissions[b]	2,002	-412	4,415

Source: Mathematica's analysis of Medicare claims data from January 2013 through December 2021.

Note: This table calculates the overall estimated effects on attributed Medicare FFS beneficiaries who were in the intent-to-treat analysis sample in Track 1 practices during the five years of CPC+. The total number of beneficiaries attributed to Track 1 practices in the annual analysis sample during the intervention period was 1,431,578. These beneficiaries had 56,551,215, eligible beneficiary months and 1,166,640 eligible index discharges (for readmissions) over the five years of CPC+. Impact estimates (shown in Tables 5.A.1.1a., 5.A.2.1a, and 5.A.3.1a) are from difference-in-differences regressions using practice fixed effects and patient-level control variables from the pre-CPC+ period. **Yellow shading *with bold, italicized text*** signifies that the estimate was statistically significant at the $p < 0.10$ level.

[a] Expenditures for Part A and B services in PY 3, PY 4, and PY 5 include QPP payment adjustments which were based on practitioner performance in 2017, 2018, and 2019 respectively. QPP payment adjustments include (1) MIPS adjustments, which were applied directly to physician and outpatient claims in 2019, 2020, and 2021 (as a percentage of the charges on the claims), and (2) lump-sum incentive payments, which were paid out to eligible practitioners who participated in Advanced APMs in 2017, 2018, and 2019; they were calculated based on applicable physician and outpatient claims for these practitioners in, respectively, 2018, 2019, and 2020. Note that the first QPP adjustments occurred in 2019 (two years after the start of QPP), so there are no QPP payments in the years before 2019.

[b] In the impact analysis, this outcome represents the percentage of discharges with an unplanned readmission within 30 days of the discharge. For this table, we translated the impact estimate into the total number of discharges for which the initiative affected readmissions.

APM = Alternative Payment Model; CI = confidence interval; CPCP = Comprehensive Primary Care Payment; CMS = Centers for Medicare and Medicaid Services; ED = emergency department; FFS = fee-for-service; MIPS = Merit-based Incentive Payment System; PY = Program Year; QPP = Quality Payment Program.

Table 5.A.4.2. Aggregate impact estimates for key outcomes across the five years of CPC+, Track 2

Outcome	Estimate	90 percent CI lower bound	90 percent CI upper bound
Medicare expenditures without CMS's enhanced payments[a]	$92,139,501	-$343,919,693	$528,198,694
Medicare expenditures including CMS's enhanced payments[a]	***$1,678,641,929***	$1,247,748,740	$2,109,535,117
Hospitalizations	***-15,766***	-30,425	-1,107
Outpatient ED visits	***-47,395***	-72,149	-22,641
30-day readmissions[b]	852	-2,172	3,876

Source: Mathematica's analysis of Medicare claims data from January 2013 through December 2021.

Note: This table calculates the overall estimated effects on attributed Medicare FFS beneficiaries who were in the intent-to-treat analysis sample in Track 2 practices during the five years of CPC+. The total number of beneficiaries attributed to Track 2 practices in the annual analysis sample during the intervention period was 1,753,421 . These beneficiaries had 68,907,660 eligible beneficiary months and 1,434,465 eligible index discharges (for readmissions) over the five years of CPC+. Impact estimates (shown in Tables 5.A.1.2a., 5.A.2.2a, and 5.A.3.2a) are from difference-in-differences regressions using practice fixed effects and patient-level control variables from the pre-CPC+ period. **Yellow shading *with bold, italicized text*** signifies that the estimate was statistically significant at the $p < 0.10$ level.

[a] Expenditures for Part A and B services in PY 3, PY 4, and PY 5 include QPP payment adjustments which were based on practitioner performance in 2017, 2018, and 2019 respectively. QPP payment adjustments include (1) MIPS adjustments, which were applied directly to physician and outpatient claims in 2019, 2020, and 2021 (as a percentage of the charges on the claims), and (2) lump-sum incentive payments, which were paid out to eligible practitioners who participated in Advanced APMs in 2017, 2018, and 2019; they were calculated based on applicable physician and outpatient claims for these practitioners in, respectively, 2018, 2019, and 2020. Note that the first QPP adjustments occurred in 2019 (two years after the start of QPP), so there are no QPP payments in the years before 2019.

[b] In the impact analysis, this outcome represents the percentage of discharges with an unplanned readmission within 30 days of the discharge. For this table, we translated the impact estimate into the total number of discharges for which the initiative affected readmissions.

APM = Alternative Payment Model; CI = confidence interval; CPCP = Comprehensive Primary Care Payment; CMS = Centers for Medicare and Medicaid Services; ED = emergency department; FFS = fee-for-service; MIPS = Merit-based Incentive Payment System; PY = Program Year; QPP = Quality Payment Program.

5.A.5. Detailed results from triple-differences sensitivity test

Table 5.A.5.1. Estimated triple-differences and difference-in-differences impacts on selected expenditures and service use outcomes for Medicare FFS beneficiaries average across the five program years, Track 1 by SSP status

		Overall			SSP			Non-SSP			
	Overall CPC+ mean	DDD percentage impact[a,b]	DD percentage impact[a,c]	*p*-Value of difference[d]	DDD percentage impact[a,b]	DD percentage impact[a,c]	*p*-Value of difference	DDD percentage impact[a,b]	DD percentage impact[a,c]	*p*-Value of difference	*p*-Value between DDD SSP and non-SSP[d]
Medicare expenditures (per beneficiary per month)											
Medicare Part A and B expenditures without enhanced payments for CPC+ and SSP[e]											
Baseline	$881	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA
PY 1	$899	0.7%	0.6%	0.66	0.8%	0.2%	0.29	0.6%	1.0%	0.46	0.83
PY 2	$949	1.1%	0.4%	0.20	1.4%	0.0%	0.07	0.6%	0.9%	0.65	0.50
PY 3	$994	0.6%	0.2%	0.44	0.5%	-0.8%	0.07	0.7%	1.4%	0.20	0.93
PY 4	$949	0.4%	-0.3%	0.17	-0.3%	-1.5%	0.11	1.4%	1.1%	0.69	0.27
PY 5	$1,042	0.5%	-0.3%	0.15	-1.0%	-1.8%	0.34	2.1%	1.0%	0.12	0.04
PY 1 through PY 5	$969	0.7%	0.1%	NA	0.3%	-0.8%	NA	1.0%	1.1%	NA	0.53
Service use (per 1,000 beneficiaries per year)											
Acute hospitalizations (short-stay acute care and critical access hospitals)											
Baseline	290	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA
PY 1	289	-0.1%	-0.2%	0.82	-0.5%	-0.9%	0.60	0.3%	0.6%	0.73	0.65
PY 2	285	0.3%	-0.6%	0.17	1.3%	-0.8%	0.08	-0.7%	-0.5%	0.80	0.27
PY 3	284	0.0%	-0.9%	0.20	-0.3%	-1.7%	0.18	0.2%	0.0%	0.82	0.83
PY 4	243	0.9%	-2.0%	0.00	0.1%	-3.2%	0.01	1.7%	-0.6%	0.02	0.50
PY 5	244	0.1%	-1.1%	0.17	-1.3%	-2.0%	0.59	1.9%	0.3%	0.11	0.15
PY 1 through PY 5	268	0.2%	-0.9%	NA	-0.1%	-1.6%	NA	0.5%	0.0%	NA	0.73
Outpatient ED visits, including observation stays											
Baseline	493	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA
PY 1	490	-0.1%	-1.1%	0.08	-0.2%	-1.2%	0.35	0.1%	-1.0%	0.09	0.77
PY 2	484	-0.5%	-1.5%	0.12	-1.1%	-1.7%	0.54	0.1%	-1.2%	0.06	0.46
PY 3	484	-0.6%	-1.6%	0.16	-1.1%	-1.6%	0.72	-0.1%	-1.7%	0.04	0.56
PY 4	376	-0.6%	-2.8%	0.04	-1.1%	-3.6%	0.15	0.1%	-1.7%	0.13	0.67
PY 5	407	-1.6%	-3.7%	0.04	-1.7%	-4.3%	0.12	-1.1%	-3.2%	0.08	0.86
PY 1 through PY 5	446	-0.6%	-2.1%	NA	-1.0%	-2.3%	NA	-0.1%	-1.8%	NA	0.60

Table 5.A.5.1. ***(continued)***

	Overall CPC+ mean	Overall			SSP			Non-SSP			*p*-Value between DDD SSP and non-SSP[d]
		DDD percentage impact[a,b]	DD percentage impact[a,c]	*p*-Value of difference[d]	DDD percentage impact[a,b]	DD percentage impact[a,c]	*p*-Value of difference	DDD percentage impact[a,b]	DD percentage impact[a,c]	*p*-Value of difference	
Total urgent care center (UCC) visits											
Baseline	104	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA
PY 1	119	0.8%	0.6%	0.87	0.3%	1.5%	0.48	1.2%	-0.5%	0.28	0.83
PY 2	135	0.3%	2.3%	0.12	-0.5%	4.7%	0.00	1.2%	-0.9%	0.23	0.73
PY 3	149	1.6%	2.4%	0.66	-1.2%	2.3%	0.21	5.7%	2.8%	0.15	0.33
PY 4	156	4.7%	15.4%	0.00	0.4%	13.8%	0.00	12.2%	19.2%	0.01	0.22
PY 5	212	0.6%	2.0%	0.55	1.2%	-1.1%	0.58	1.6%	5.7%	0.08	0.99
PY 1 through PY 5	156	1.5%	4.2%	NA	0.1%	3.7%	NA	4.0%	5.0%	NA	0.49
Ambulatory primary care visits (including FQHCs, RHCs, and CAHs)[f]											
Baseline	4,255	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA
PY 1	4,295	-0.9%	-1.2%	0.30	-0.4%	-1.0%	0.21	-1.5%	-1.5%	0.97	0.25
PY 2	4,340	0.2%	-0.4%	0.16	0.6%	-0.1%	0.34	-0.2%	-0.8%	0.29	0.49
PY 3	4,406	1.1%	0.0%	0.03	1.4%	0.1%	0.12	0.9%	-0.1%	0.12	0.76
PY 4	3,991	0.1%	-0.5%	0.44	-0.2%	-0.1%	0.93	0.3%	-0.8%	0.20	0.78
PY 5	4,244	1.1%	0.0%	0.09	1.5%	-0.1%	0.12	0.5%	-0.2%	0.38	0.64
PY 1 through PY 5	4,252	0.3%	-0.4%	NA	0.5%	-0.2%	NA	0.0%	-0.6%	NA	0.64
Proportion of ambulatory primary care visits that are not face-to-face											
PY 4	15.7%	-0.4%	5.9%	0.00	-8.1%	4.7%	0.00	9.7%	11.9%	0.23	0.14
PY 5	8.4%	-2.6%	4.3%	0.00	-11.4%	1.4%	0.00	8.6%	10.3%	0.26	0.15

Source: Mathematica's analysis of Medicare claims data from January 2013 through December 2021.

Notes: Although this table indicates which estimates are statistically significant, when we interpret evidence, we combine evidence from the magnitude of the effect, the *p*-values, findings on related outcomes, and sensitivity tests.

[a] We calculated percentage impacts relative to what the CPC+ mean would have been in each year in the absence of the intervention—that is, the unadjusted CPC+ mean minus the impact estimate.

[b] Triple-differences impact estimates are regression-adjusted for pre-CPC+ beneficiary characteristics (including HCC scores), practice fixed effects, and COVID-19 controls. Each impact estimate except for the outcome for proportion of ambulatory primary care visits that were not face-to-face is based on a triple-differences analysis and reflects the difference between (1) the difference in the regression-adjusted average outcome for attributed Medicare FFS beneficiaries in CPC+ practices in the five program years compared with baseline relative to the same difference over time for attributed Medicare FFS beneficiaries in comparison practices, and (2) the difference in the regression-adjusted average outcome for attributed Medicare FFS beneficiaries in non-CPC+ practices in the five program years compared with baseline relative to the same difference over time for attributed Medicare FFS beneficiaries in non-comparison practices. Impact estimates for the proportion of ambulatory primary care visits that were not face-to-face reflect the difference between (1) the difference in the regression-adjusted average outcome for attributed Medicare FFS beneficiaries in CPC+ practices and comparison practices in PY 4 or 5, and (2) the difference in the regression-adjusted average outcome for attributed Medicare FFS beneficiaries in non-CPC+ practices and non-comparison practices in PY 4 or 5.

[c] Difference-in-differences impact estimates are regression-adjusted using a difference-in-differences analysis that reflects the difference between the average outcome for Medicare FFS beneficiaries attributed to CPC+ practices in the five years of CPC+ and the average outcome in the baseline year, relative to the same difference over time for Medicare FFS beneficiaries attributed to comparison practices, while controlling for beneficiary characteristics, practice fixed effects, and COVID-19 controls. Impact estimates for the proportion of ambulatory primary care visits that are not face-to-face reflect the difference between the regression-adjusted average outcome for attributed Medicare FFS beneficiaries in CPC+ practices and comparison practices in PY 4 and PY 5.

[d] Represents the *p*-value of the underlying impact estimate (which is in dollars PBPM for expenditures, per 1,000 beneficiaries per year for continuous measures of service use, and in percentage points for binary measures of service use).

Table 5.A.5.1. ***(continued)***

[e] Expenditures for Part A and Part B services in PY 3 through PY 5 include QPP payment adjustments, based on practitioner performance two years before. They are applicable for CPC+, comparison, non-CPC+, and non-comparison practices. The adjustments are composed of (1) MIPS adjustments, which are applied directly to physician and outpatient claims (as a percentage of the charges on the claims); and (2) lump sum incentive payments to eligible practitioners who participated in Advanced APMs in 2017, 2018, and 2019 (calculated based on 2018, 2019, and 2020 claims for these practitioners, respectively). The first QPP adjustments were paid in PY 3 (two years after the start of QPP), so there are no QPP payments in PYs 1 and 2.

[f] Ambulatory visits with primary care practitioners and specialists include office-based visits and visits at home, as well as visits in other settings, such as FQHCs, RHCs, and CAHs.

*/**/*** Underlying impact estimate (which is in dollars PBPM for expenditures, per 1,000 beneficiaries per year for measures of service use) is significantly different from zero at the 0.10/0.05/0.01 level, two-tailed test.

APM = Alternative Payment Model; CAH = critical access hospital; DD = difference-in-differences; DDD = triple-differences; ED = emergency department; FFS = fee-for-service; FQHC = Federally Qualified Health Center; NA = not applicable; SSP = Medicare Shared Savings Program; RHC = Rural Health Clinic; PBPM = per beneficiary per month; PY = Program Year; QPP = Quality Payment Program.

Table 5.A.5.2. Estimated triple-differences and difference-in-differences impacts on selected expenditures and service use outcomes for Medicare FFS beneficiaries average across the five program years, Track 2 by SSP status

		Overall			SSP			Non-SSP			
	Overall CPC+ mean	DDD percentage impact[a,b]	DD percentage impact[a,c]	*p*-Value of difference[d]	DDD percentage impact[a,b]	DD percentage impact[a,c]	*p*-Value of difference	DDD percentage impact[a,b]	DD percentage impact[a,c]	*p*-Value of difference	*p*-Value between DDD SSP and non-SSP[d]
Medicare expenditures (per beneficiary per month)											
Medicare Part A and B expenditures without enhanced payments for CPC+ and SSP[e]											
Baseline	$876	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA
PY 1	$897	0.6%	0.6%	0.90	0.5%	0.1%	0.65	0.6%	0.9%	0.54	0.99
PY 2	$949	0.6%	0.5%	0.84	0.0%	-0.3%	0.84	1.1%	1.2%	0.90	0.47
PY 3	$989	0.3%	-0.2%	0.37	0.3%	-0.8%	0.20	0.1%	0.3%	0.69	0.87
PY 4	$946	0.4%	-0.2%	0.30	-0.3%	-1.5%	0.23	1.0%	1.0%	0.90	0.41
PY 5	$1,034	1.0%	0.1%	0.10	-0.8%	-1.6%	0.41	2.2%	1.4%	0.20	0.07
PY 1 through PY 5	$965	0.6%	0.1%	NA	0.0%	-0.8%	NA	1.0%	0.9%	NA	0.42
Service use (per 1,000 beneficiaries per year)											
Acute hospitalizations (short-stay acute care and critical access hospitals)											
Baseline	292	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA
PY 1	292	-0.1%	-0.2%	0.86	0.0%	-0.1%	0.85	-0.3%	-0.2%	0.93	0.83
PY 2	289	-0.2%	-0.5%	0.72	-0.2%	0.0%	0.91	-0.4%	-0.9%	0.46	0.92
PY 3	286	-0.8%	-1.7%	0.23	0.5%	-0.7%	0.33	-2.1%	-2.5%	0.64	0.18
PY 4	245	0.2%	-1.9%	0.01	1.2%	-1.5%	0.04	-0.8%	-2.1%	0.16	0.36
PY 5	246	-0.9%	-0.8%	0.83	-2.2%	-0.9%	0.33	-0.2%	-0.8%	0.54	0.37
PY 1 through PY 5	270	-0.4%	-1.0%	NA	-0.1%	-0.6%	NA	-0.8%	-1.3%	NA	0.66
Outpatient ED visits, including observation stays											
Baseline	492	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA
PY 1	486	-0.2%	-1.6%	0.04	-0.2%	-1.9%	0.20	-0.2%	-1.3%	0.10	0.96
PY 2	483	-0.7%	-1.4%	0.42	-1.5%	-1.7%	0.89	-0.1%	-1.1%	0.20	0.50
PY 3	483	-1.0%	-1.5%	0.52	-2.2%	-1.6%	0.73	0.0%	-1.4%	0.07	0.30
PY 4	378	1.7%	-2.2%	0.00	-1.0%	-5.0%	0.02	4.1%	0.6%	0.00	0.07
PY 5	408	0.9%	-2.7%	0.00	-0.6%	-4.9%	0.04	2.6%	-0.3%	0.02	0.28
PY 1 through PY 5	445	0.0%	-1.8%	NA	-1.2%	-2.9%	NA	0.9%	-0.8%	NA	0.27

Table 5.A.5.2. ***(continued)***

		Overall			SSP			Non-SSP			
	Overall CPC+ mean	DDD percentage impact[a,b]	DD percentage impact[a,c]	*p*-Value of difference[d]	DDD percentage impact[a,b]	DD percentage impact[a,c]	*p*-Value of difference	DDD percentage impact[a,b]	DD percentage impact[a,c]	*p*-Value of difference	*p*-Value between DDD SSP and non-SSP[d]
Total urgent care center (UCC) visits											
Baseline	97	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA
PY 1	111	1.7%	1.0%	0.67	0.9%	3.4%	0.39	2.1%	-0.8%	0.06	0.83
PY 2	124	1.6%	1.8%	0.90	-1.1%	6.9%	0.04	3.6%	-2.3%	0.00	0.49
PY 3	134	1.0%	-1.7%	0.33	-2.7%	-0.1%	0.61	3.6%	-3.0%	0.00	0.44
PY 4	136	2.6%	7.1%	0.08	-1.5%	7.8%	0.02	6.5%	6.2%	0.93	0.31
PY 5	186	1.4%	3.1%	0.59	0.0%	-2.1%	0.74	1.3%	5.8%	0.07	0.89
PY 1 through PY 5	140	1.6%	2.2%	NA	-0.9%	2.5%	NA	3.2%	1.2%	NA	0.50
Ambulatory primary care visits (including FQHCs, RHCs, and CAHs)[f]											
Baseline	4,361	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA
PY 1	4,364	-1.0%	-1.6%	0.06	0.2%	-1.1%	0.03	-1.8%	-2.0%	0.70	0.04
PY 2	4,393	-0.8%	-1.0%	0.56	-0.7%	-0.6%	0.94	-0.8%	-1.3%	0.33	0.90
PY 3	4,449	-0.3%	-0.8%	0.40	-0.4%	-0.4%	1.00	-0.3%	-1.1%	0.20	0.95
PY 4	4,019	-1.2%	-0.7%	0.66	-1.8%	0.0%	0.33	-0.4%	-1.3%	0.27	0.56
PY 5	4,236	1.5%	-1.0%	0.07	2.8%	-0.2%	0.17	-0.6%	-1.5%	0.25	0.23
PY 1 through PY 5	4,286	-0.4%	-1.0%	NA	0.0%	-0.5%	NA	-0.8%	-1.4%	NA	0.54
Proportion of ambulatory primary care visits that are not face-to-face											
PY 4	16.9%	3.9%	14.6%	0.00	-2.0%	11.7%	0.00	5.4%	13.1%	0.00	0.82
PY 5	8.9%	2.1%	13.8%	0.00	-0.8%	14.5%	0.00	4.5%	12.7%	0.02	0.60

Source: Mathematica's analysis of Medicare claims data from January 2013 through December 2021.

Notes: Although this table indicates which estimates are statistically significant, when we interpret evidence, we combine evidence from the magnitude of the effect, the *p*-values, findings on related outcomes, and sensitivity tests.

[a] We calculated percentage impacts relative to what the CPC+ mean would have been in each year in the absence of the intervention—that is, the unadjusted CPC+ mean minus the impact estimate.

[b] Triple-differences impact estimates are regression-adjusted for pre-CPC+ beneficiary characteristics (including HCC scores), practice fixed effects, and COVID-19 controls. Each impact estimate except for the outcome for proportion of ambulatory primary care visits that were not face-to-face is based on a triple-differences analysis and reflects the difference between (1) the difference in the regression-adjusted average outcome for attributed Medicare FFS beneficiaries in CPC+ practices in the five program years compared with baseline relative to the same difference over time for attributed Medicare FFS beneficiaries in comparison practices, and (2) the difference in the regression-adjusted average outcome for attributed Medicare FFS beneficiaries in non-CPC+ practices in the five program years compared with baseline relative to the same difference over time for attributed Medicare FFS beneficiaries in non-comparison practices. Impact estimates for the proportion of ambulatory primary care visits that were not face-to-face reflect the difference between (1) the difference in the regression-adjusted average outcome for attributed Medicare FFS beneficiaries in CPC+ practices and comparison practices in PY 4 or 5, and (2) the difference in the regression-adjusted average outcome for attributed Medicare FFS beneficiaries in non-CPC+ practices and non-comparison practices in PY 4 or 5.

[c] Difference-in-differences impact estimates are regression-adjusted using a difference-in-differences analysis that reflects the difference between the average outcome for Medicare FFS beneficiaries attributed to CPC+ practices in the five years of CPC+ and the average outcome in the baseline year, relative to the same difference over time for Medicare FFS beneficiaries attributed to comparison practices, while controlling for beneficiary characteristics, practice fixed effects, and COVID-19 controls. Impact estimates for the proportion of ambulatory primary care visits that are not face-to-face reflect the difference between the regression-adjusted average outcome for attributed Medicare FFS beneficiaries in CPC+ practices and comparison practices in PY 4 and PY 5.

[d] Represents the *p*-value of the underlying impact estimate (which is in dollars PBPM for expenditures, per 1,000 beneficiaries per year for continuous measures of service use, and in percentage points for binary measures of service use).

Table 5.A.5.2. ***(continued)***

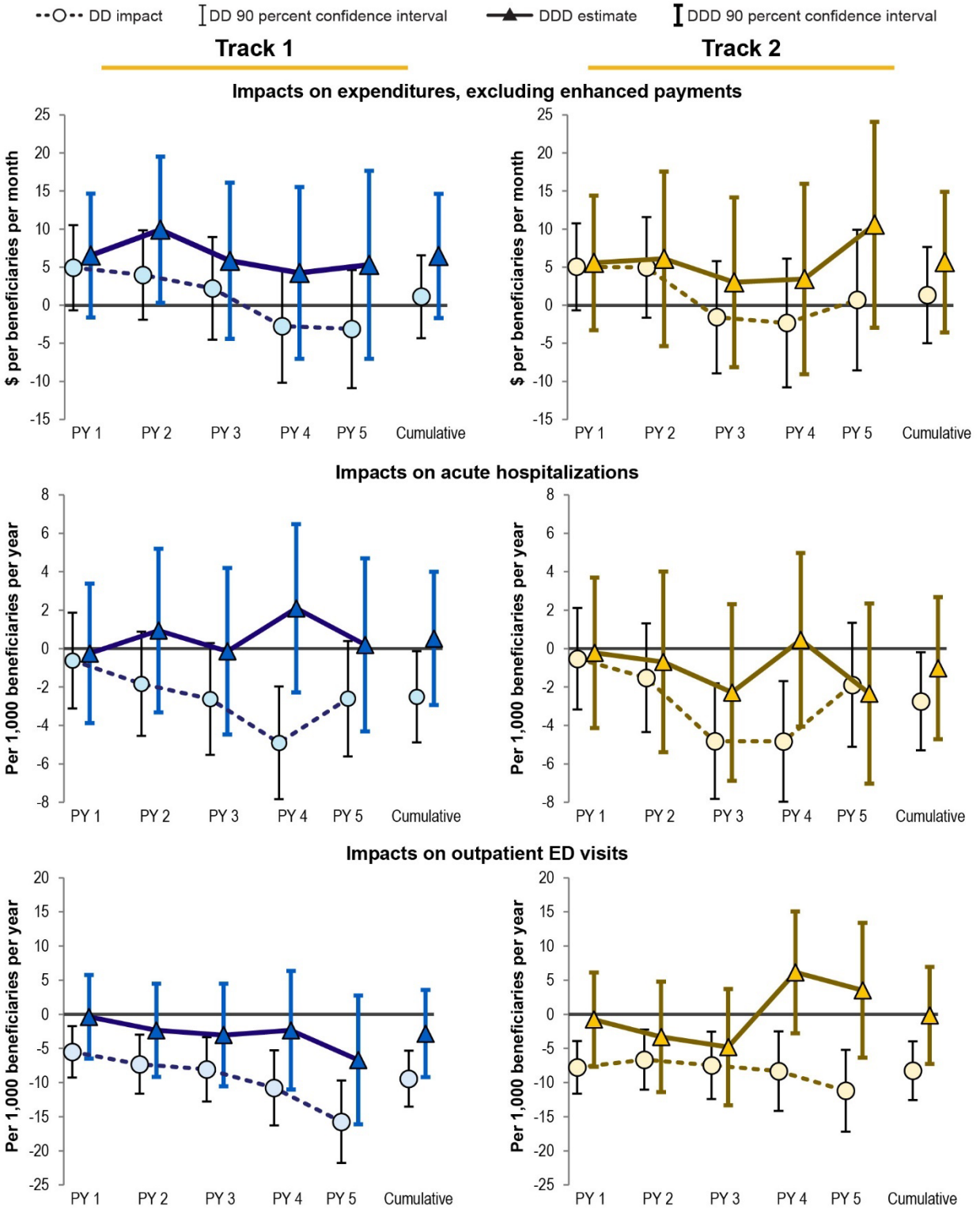
[e] Expenditures for Part A and Part B services in PY 3 through PY 5 include QPP payment adjustments, based on practitioner performance two years before. They are applicable for CPC+, comparison, non-CPC+, and non-comparison practices. The adjustments are composed of (1) MIPS adjustments, which are applied directly to physician and outpatient claims (as a percentage of the charges on the claims); and (2) lump sum incentive payments to eligible practitioners who participated in Advanced APMs in 2017, 2018, and 2019 (calculated based on 2018, 2019, and 2020 claims for these practitioners, respectively). The first QPP adjustments were paid in PY 3 (two years after the start of QPP), so there are no QPP payments in PYs 1 and 2.

[f] Ambulatory visits with primary care practitioners and specialists include office-based visits and visits at home, as well as visits in other settings, such as FQHCs, RHCs, and CAHs.

*/**/*** Underlying impact estimate (which is in dollars PBPM for expenditures, per 1,000 beneficiaries per year for measures of service use) is significantly different from zero at the 0.10/0.05/0.01 level, two-tailed test.

APM = Alternative Payment Model; CAH = critical access hospital; DD = difference-in-differences; DDD = triple-differences; ED = emergency department; FFS = fee-for-service; FQHC = Federally Qualified Health Center; NA = not applicable; SSP = Medicare Shared Savings Program; RHC = Rural Health Clinic; PBPM = per beneficiary per month; PY = Program Year; QPP = Quality Payment Program.

Figure 5.A.5.1. Triple-differences and difference-in-differences model impact estimates for expenditures, acute hospitalizations, outpatient ED visits, urgent care center visits, and ambulatory primary care visits by program year and average across the five program years, Tracks 1 and 2



Source: Mathematica's analyses of Medicare claims data from January 2013 through December 2021.

Figure 5.A.5.1. ***(continued)***

[a] Triple-differences impact estimates are regression-adjusted for pre-CPC+ beneficiary characteristics (including HCC scores), practice fixed effects, and COVID-19 controls. Each impact estimate is based on a triple-differences analysis and reflects the difference between (1) the difference in the regression-adjusted average outcome for attributed Medicare FFS beneficiaries in CPC+ practices in the five program years compared with baseline relative to the same difference over time for attributed Medicare FFS beneficiaries in comparison practices, and (2) the difference in the regression-adjusted average outcome for attributed Medicare FFS beneficiaries in non-CPC+ practices in the five program years compared with baseline relative to the same difference over time for attributed Medicare FFS beneficiaries in non-comparison practices.

[b] Difference-in-differences impact estimates are regression-adjusted using a difference-in-differences analysis that reflects the difference between the average outcome for Medicare FFS beneficiaries attributed to CPC+ practices in the five years of CPC+ and the average outcome in the baseline year, relative to the same difference over time for Medicare FFS beneficiaries attributed to comparison practices, while controlling for beneficiary characteristics, practice fixed effects, and COVID-19 controls.

CI = confidence interval; DD = difference-in-differences; DDD = triple-differences; ED = emergency department; FFS = fee-for-service; HCC = hierarchical condition category; PCP = primary care practitioner; PY = Program Year.

Table 5.A.5.3a. Regression-adjusted means and estimated triple-differences impacts of CPC+ on selected expenditures and service use outcomes for attributed Medicare FFS beneficiaries by program year and average across the five program years, Track 1

	Overall					
	CPC+ mean[a]	Comparison mean[a]	Non-CPC+ mean[a]	Non-comparison mean[a]	Impact estimate[b] (SE)	90% confidence interval
Medicare expenditures (per beneficiary per month)						
Total Medicare Part A and B expenditures without enhanced payments for CPC+ and SSP						
Baseline	$881	$884	$938	$936	NA	NA
PY 1	$899	$897	$954	$957	$6.5 ($5.0)	(-$1.6, $14.7)
PY 2	$949	$949	$1,010	$1,017	$9.9* ($5.8)	($0.3, $19.5)
PY 3	$994	$997	$1,048	$1,054	$5.8 ($6.2)	(-$4.4, $16.1)
PY 4	$949	$962	$996	$1,012	$4.2 ($6.8)	(-$7.0, $15.5)
PY 5	$1,042	$1,057	$1,073	$1,092	$5.3 ($7.5)	(-$7.1, $17.6)
PY 1 through PY 5	$969	$976	$1,014	$1,025	$6.5 ($5.0)	(-$1.7, $14.6)
Service use (per 1,000 beneficiaries per year)						
Acute hospitalizations (short-stay acute care and critical access hospitals)						
Baseline	290	289	320	305	NA	NA
PY 1	289	287	318	305	-0.3 (2.2)	(-3.9, 3.4)
PY 2	285	285	318	307	0.9 (2.6)	(-3.3, 5.2)
PY 3	284	285	313	302	-0.1 (2.6)	(-4.5, 4.2)
PY 4	243	246	264	259	2.1 (2.7)	(-2.3, 6.5)
PY 5	244	247	262	254	0.2 (2.7)	(-4.3, 4.7)
PY 1 through PY 5	268	269	296	286	0.5 (2.1)	(-2.9, 4.0)
Outpatient ED visits, including observation stays						
Baseline	493	498	547	547	NA	NA
PY 1	490	497	543	547	-0.4 (3.7)	(-6.5, 5.8)
PY 2	484	495	527	534	-2.3 (4.2)	(-9.2, 4.5)
PY 3	484	493	521	527	-3.0 (4.6)	(-10.5, 4.5)
PY 4	376	388	402	415	-2.3 (5.3)	(-11.0, 6.4)
PY 5	407	421	431	444	-6.7 (5.7)	(-16.1, 2.7)
PY 1 through PY 5	446	456	487	494	-2.8 (3.9)	(-9.2, 3.6)

Table 5.A.5.3a. ***(continued)***

	Overall					
	CPC+ mean[a]	Comparison mean[a]	Non-CPC+ mean[a]	Non-comparison mean[a]	Impact estimate[b] (SE)	90% confidence interval
Total urgent care center (UCC) visits						
Baseline	104	111	92	99	NA	NA
PY 1	119	124	104	111	0.9 (2.2)	(-2.7, 4.5)
PY 2	135	138	111	114	0.4 (3.1)	(-4.7, 5.5)
PY 3	149	152	124	128	2.4 (4.6)	(-5.2, 10.0)
PY 4	156	144	134	125	7.1 (5.5)	(-2.0, 16.2)
PY 5	212	202	186	171	1.2 (8.1)	(-12.1, 14.4)
PY 1 through PY 5	156	154	130	129	2.3 (3.7)	(-3.8, 8.3)
Ambulatory primary care visits						
Baseline	4,255	4,370	4,586	4,628	NA	NA
PY 1	4,295	4,466	4,666	4,723	-40.0** (20.4)	(-73.6, -6.4)
PY 2	4,340	4,480	4,669	4,745	8.8 (27.2)	(-35.8, 53.5)
PY 3	4,406	4,519	4,693	4,780	49.3 (31.4)	(-2.5, 101.0)
PY 4	3,991	4,137	4,301	4,375	2.7 (39.7)	(-62.5, 68.0)
PY 5	4,244	4,380	4,470	4,589	46.6 (41.7)	(-22.0, 115.3)
PY 1 through PY 5	4,252	4,390	4,565	4,644	12.4 (26.7)	(-31.5, 56.3)
Proportion of ambulatory primary care visits that were not face-to-face						
PY 4	15.7%	15.2%	15.1%	14.4%	-0.1 (0.4)	(-0.8, 0.7)
PY 5	8.4%	7.9%	8.5%	7.8%	-0.2 (0.3)	(-0.7, 0.2)
Unweighted sample sizes						
Number of practices	1,373	5,243	8,337	20,656		
Number of beneficiaries	1,549,585	5,347,499	4,015,775	11,444,943		
Number of beneficiary years[c]	5,916,394	20,150,090	14,995,442	43,307,169		

Source: Mathematica's analysis of Medicare claims data from January 2013 through December 2021.

Notes: Although this table indicates which estimates are statistically significant, when we interpret evidence, we combine evidence from the magnitude of the effect, the *p*-values, findings on related outcomes, and sensitivity tests.

[a] We report the actual, unadjusted CPC+ mean for each time period shown in the table. For comparison group practices, non-CPC+ practices, and non-comparison practices, we report the actual, unadjusted mean during the baseline period but the adjusted mean during each intervention period. We obtained the adjusted mean by subtracting the regression-adjusted difference between the CPC+ mean and each group's mean in each time period from the CPC+ mean in that same time period.

Table 5.A.5.3a. ***(continued)***

[b] Impact estimates are regression-adjusted for pre-CPC+ beneficiary characteristics (including HCC scores), practice fixed effects, and COVID-19 controls. Each impact estimate except for the non-face-to-face percentage of ambulatory primary care visits outcome is based on a triple-differences analysis and reflects the difference between (1) the difference in the regression-adjusted average outcome for attributed Medicare FFS beneficiaries in CPC+ practices in the five program years compared with baseline relative to the same difference over time for attributed Medicare FFS beneficiaries in comparison practices, and (2) the difference in the regression-adjusted average outcome for attributed Medicare FFS beneficiaries in non-CPC+ practices in the five program years compared with baseline relative to the same difference over time for attributed Medicare FFS beneficiaries in non-comparison practices. Impact estimates for the non-face-to-face percentage of ambulatory primary care visits outcome reflect the difference between (1) the difference in the regression-adjusted average outcome for attributed Medicare FFS beneficiaries in CPC+ practices and comparison practices in PY 4 (or PY 5), and (2) the difference in the regression-adjusted average outcome for attributed Medicare FFS beneficiaries in non-CPC+ practices and non-comparison practices in PY 4 (or PY 5).

[c] After accounting for weights that adjust for matching, time observed in Medicare FFS, and the concentration of CPC+ in each geographic area, the effective sample sizes are reduced. For non-CPC+ practices, the effective sample size (in terms of beneficiary-years) is 30 percent of the actual group size. For non-comparison practices, the effective sample size (in terms of beneficiary-years) is 16 percent of the actual group size. For the comparison group, the effective sample size (in terms of beneficiary-years) is 45 percent of the size of the actual comparison group. Because the CPC+ sample size is affected only by time the beneficiary is observed (and is not affected by the matching weights), the effective sample size for the CPC+ group is about 96 percent of the actual sample size.

*/**/*** Significantly different from zero at the 0.10/0.05/0.01 level, two-tailed test.

ED = emergency department; FFS = fee-for-service; HCC = hierarchical condition category; NA = not applicable; non-comparison = unselected practices in comparison regions; non-CPC+ = non-participating practices in CPC+ regions; PY = Program Year; SE = standard error.

Table 5.A.5.3b. Regression-adjusted means and estimated triple-differences impacts of CPC+ on selected expenditures and service use outcomes for attributed Medicare FFS beneficiaries by program year and average across the five program years, Track 1 by SSP status

	Track 1 – SSP						Track 1 – Non-SSP					
	CPC+ mean[a]	Comparison mean[a]	Non-CPC+ mean[a]	Non-comparison mean[a]	Impact estimate[b] (SE)	90% confidence interval	CPC+ mean[a]	Comparison mean[a]	Non-CPC+ mean[a]	Non-comparison mean[a]	Impact estimate[b] (SE)	90% confidence interval
Medicare expenditures (per beneficiary per month)												
Total Medicare Part A and B expenditures without enhanced payments for CPC+ and SSP												
Baseline	906	905	966	954	NA	NA	854	861	909	915	NA	NA
PY 1	924	918	978	979	$7.3 ($7.2)	(-$4.5, $19.1)	874	873	927	931	$5.3 ($6.7)	(-$5.7, $16.4)
PY 2	974	973	1,033	1,038	$13.3 ($8.9)	(-$1.3, $28.0)	923	922	985	992	$5.8 ($7.4)	(-$6.4, $18.0)
PY 3	1,017	1,025	1,069	1,074	$5.1 ($9.1)	(-$9.9, $20.1)	971	966	1,025	1,032	$6.4 ($8.4)	(-$7.4, $20.1)
PY 4	969	1,025	1,017	1,022	-$2.6 ($10.2)	(-$19.3, $14.1)	929	938	974	1,001	$12.5 ($8.9)	(-$2.2, $27.2)
PY 5	1,073	1,025	1,102	1,101	-$10.7 ($11.3)	(-$29.2, $7.8)	1,009	1,026	1,041	1,082	$20.7** ($9.8)	($4.6, $36.8)
PY 1 through PY 5	994	1,025	1,038	1,041	$3.4 ($7.3)	(-$8.6, $15.5)	944	949	988	1,007	$9.8 ($6.6)	(-$1.0, $20.6)
Service use (per 1,000 beneficiaries per year)												
Acute hospitalizations (short-stay acute care and critical access hospitals)												
Baseline	291	289	321	304	NA	NA	289	288	318	306	NA	NA
PY 1	289	288	320	305	-1.4 (3.1)	(-6.4, 3.7)	289	286	317	304	0.9 (3.1)	(-4.3, 6.1)
PY 2	286	287	318	307	3.5 (3.9)	(-2.8, 9.9)	283	283	317	306	-2.0 (3.4)	(-7.5, 3.6)
PY 3	286	289	316	302	-0.7 (3.8)	(-7.0, 5.5)	283	281	310	301	0.5 (3.6)	(-5.4, 6.4)
PY 4	245	249	268	259	0.3 (3.9)	(-6.1, 6.6)	241	242	260	258	4.0 (3.6)	(-2.0, 10.0)
PY 5	250	253	268	253	-3.4 (4.1)	(-10.1, 3.3)	239	240	256	256	4.5 (3.6)	(-1.4, 10.5)
PY 1 through PY 5	270	272	299	286	-0.2 (3.0)	(-5.2, 4.8)	266	265	293	286	1.4 (2.9)	(-3.3, 6.1)

Table 5.A.5.3b. ***(continued)***

	Track 1 – SSP						Track 1 – Non-SSP					
	CPC+ mean[a]	Comparison mean[a]	Non-CPC+ mean[a]	Non-comparison mean[a]	Impact estimate[b] (SE)	90% confidence interval	CPC+ mean[a]	Comparison mean[a]	Non-CPC+ mean[a]	Non-comparison mean[a]	Impact estimate[b] (SE)	90% confidence interval
Outpatient ED visits, including observation stays												
Baseline	476	480	527	529	NA	NA	510	518	567	568	NA	NA
PY 1	475	478	528	529	-1.2 (5.6)	(-10.3, 8.0)	506	519	559	567	0.5 (4.8)	(-7.4, 8.4)
PY 2	467	476	512	514	-5.1 (6.2)	(-15.3, 5.0)	502	516	543	556	0.7 (5.5)	(-8.3, 9.8)
PY 3	469	474	507	506	-5.4 (6.8)	(-16.7, 5.8)	499	513	537	551	-0.5 (6.0)	(-10.3, 9.3)
PY 4	361	372	387	398	-4.0 (8.0)	(-17.2, 9.2)	392	406	419	434	0.2 (6.9)	(-11.1, 11.5)
PY 5	395	406	418	427	-6.6 (8.5)	(-20.6, 7.3)	419	437	444	463	-4.8 (7.5)	(-17.2, 7.5)
PY 1 through PY 5	431	438	473	476	-4.3 (5.8)	(-13.9, 5.2)	461	475	503	515	-0.7 (5.0)	(-9.0, 7.6)
Total urgent care center (UCC) visits												
Baseline	114	112	94	104	NA	NA	93	109	90	92	NA	NA
PY 1	132	127	108	116	0.4 (3.1)	(-4.6, 5.5)	105	121	99	105	1.3 (3.1)	(-3.7, 6.3)
PY 2	151	140	119	120	-0.8 (3.9)	(-7.2, 5.6)	118	135	103	108	1.4 (4.9)	(-6.6, 9.4)
PY 3	167	159	134	137	-2.0 (6.4)	(-12.4, 8.5)	131	145	112	118	7.0 (6.7)	(-4.0, 18.0)
PY 4	179	152	149	133	0.7 (7.0)	(-10.8, 12.3)	133	136	117	115	14.5* (8.8)	(0.0, 28.9)
PY 5	251	219	211	190	2.9 (13.7)	(-19.7, 25.4)	173	182	159	150	2.7 (9.2)	(-12.5, 17.8)
PY 1 through PY 5	178	162	143	139	0.1 (5.0)	(-8.1, 8.3)	133	145	117	119	5.1 (5.3)	(-3.6, 13.9)

Table 5.A.5.3b. ***(continued)***

	Track 1 – SSP						Track 1 – Non-SSP					
	CPC+ mean[a]	Comparison mean[a]	Non-CPC+ mean[a]	Non-comparison mean[a]	Impact estimate[b] (SE)	90% confidence interval	CPC+ mean[a]	Comparison mean[a]	Non-CPC+ mean[a]	Non-comparison mean[a]	Impact estimate[b] (SE)	90% confidence interval
Ambulatory primary care visits												
Baseline	4,207	4,341	4,538	4,508	NA	NA	4,305	4,403	4,637	4,767	NA	NA
PY 1	4,260	4,337	4,617	4,620	-18.2 (28.0)	(-64.2, 27.8)	4,332	4,498	4,719	4,843	-64.5** (29.6)	(-113.2, -15.9)
PY 2	4,297	4,444	4,636	4,653	26.4 (38.8)	(-37.4, 90.2)	4,386	4,519	4,706	4,849	-10.6 (37.7)	(-72.5, 51.4)
PY 3	4,362	4,490	4,668	4,689	58.3 (44.4)	(-14.7, 131.2)	4,451	4,552	4,720	4,884	39.5 (44.3)	(-33.4, 112.3)
PY 4	3,956	4,102	4,325	4,297	-9.7 (55.1)	(-100.4, 81.0)	4,026	4,176	4,276	4,465	12.2 (55.7)	(-79.5, 103.8)
PY 5	4,218	4,347	4,475	4,517	62.1 (57.6)	(-32.6, 156.8)	4,271	4,417	4,464	4,673	23.3 (60.4)	(-76.1, 122.7)
PY 1 through PY 5	4,216	4,358	4,548	4,556	22.7 (37.0)	(-38.2, 83.5)	4,289	4,426	4,583	4,745	-1.3 (38.0)	(-63.9, 61.2)
Proportion of ambulatory primary care visits that were not face-to-face												
PY 4	16.0%	16.0%	17.0%	15.0%	-1.4** (0.6)	(-2.4, -0.4)	15.0%	14.0%	13.0%	13.0%	1.3** (0.6)	(0.3, 2.4)
PY 5	9.0%	8.0%	9.0%	8.0%	-1.1*** (0.4)	(-1.8, -0.4)	8.0%	8.0%	8.0%	7.0%	0.6* (0.4)	(0.1, 1.2)
Unweighted sample sizes												
Number of practices	738	2,979	2,488	5,151			635	2,264	5,849	15,505		
Number of beneficiaries	798,817	3,129,830	1,454,371	3,739,960			753,337	2,233,041	2,593,004	7,811,130		
Number of beneficiary years[c]	3,017,546	11,762,356	5,310,838	13,709,969			2,898,848	8,387,734	9,684,604	29,597,200		

Source: Mathematica's analysis of Medicare claims data from January 2013 through December 2021.

Notes: Although this table indicates which estimates are statistically significant, when we interpret evidence, we combine evidence from the magnitude of the effect, the *p*-values, findings on related outcomes, and sensitivity tests.

[a] We report the actual, unadjusted CPC+ mean for each time period shown in the table. For comparison group practices, non-CPC+ practices, and non-comparison practices, we report the actual, unadjusted mean during the baseline period but the adjusted mean during each intervention period. We obtained the adjusted mean by subtracting the regression-adjusted difference between the CPC+ mean and each group's mean in each time period from the CPC+ mean in that same time period.

Table 5.A.5.3b. ***(continued)***

[b] Impact estimates are regression-adjusted for pre-CPC+ beneficiary characteristics (including HCC scores), practice fixed effects, and COVID-19 controls. Each impact estimate except for the non-face-to-face percentage of ambulatory primary care visits outcome is based on a triple-differences analysis and reflects the difference between (1) the difference in the regression-adjusted average outcome for attributed Medicare FFS beneficiaries in CPC+ practices in the five program years compared with baseline relative to the same difference over time for attributed Medicare FFS beneficiaries in comparison practices, and (2) the difference in the regression-adjusted average outcome for attributed Medicare FFS beneficiaries in non-CPC+ practices in the five program years compared with baseline relative to the same difference over time for attributed Medicare FFS beneficiaries in non-comparison practices. Impact estimates for the non-face-to-face percentage of ambulatory primary care visits outcome reflect the difference between (1) the difference in the regression-adjusted average outcome for attributed Medicare FFS beneficiaries in CPC+ practices and comparison practices in PY 4 (or PY 5), and (2) the difference in the regression-adjusted average outcome for attributed Medicare FFS beneficiaries in non-CPC+ practices and non-comparison practices in PY 4 (or PY 5).

[c] After accounting for weights that adjust for matching, time observed in Medicare FFS, and the concentration of CPC+ in each geographic area, the effective sample sizes are reduced. For the SSP group: for non-CPC+ practices, the effective sample size (in terms of beneficiary-years) is 40 percent of the actual group size. For non-comparison practices, the effective sample size (in terms of beneficiary-years) is 18 percent of the actual group size. For the comparison group, the effective sample size (in terms of beneficiary-years) is 50 percent of the size of the actual comparison group. Because the CPC+ sample size is affected only by time the beneficiary is observed (and is not affected by the matching weights), the effective sample size for the CPC+ group is about 96 percent of the actual sample size. For the non-SSP group: for non-CPC+ practices, the effective sample size (in terms of beneficiary-years) is 25 percent of the actual group size. For non-comparison practices, the effective sample size (in terms of beneficiary-years) is 24 percent of the actual group size. For the comparison group, the effective sample size (in terms of beneficiary-years) is 43 percent of the size of the actual comparison group. Because the CPC+ sample size is affected only by time the beneficiary is observed (and is not affected by the matching weights), the effective sample size for the CPC+ group is about 96 percent of the actual sample size.

*/**/*** Significantly different from zero at the 0.10/0.05/0.01 level, two-tailed test.

ED = emergency department; FFS = fee-for-service; HCC = hierarchical condition category; NA = not applicable; non-comparison = unselected practices in comparison regions; non-CPC+ = non-participating practices in CPC+ regions; PY = Program Year; SE = standard error; SSP = Medicare Shared Savings Program.

Table 5.A.5.4a. Regression-adjusted means and estimated triple-differences impacts of CPC+ on selected expenditures and service use outcomes for attributed Medicare FFS beneficiaries by program year and average across the five program years, Track 2

	Track 2 – Overall					
	CPC+ mean[a]	Comparison mean[a]	Non-CPC+ mean[a]	Non-comparison mean[a]	Impact estimate[b] (SE)	90% confidence interval
Medicare expenditures (per beneficiary per month)						
Total Medicare Part A and B expenditures without enhanced payments for CPC+ and SSP						
Baseline	$876	$877	$928	$931	NA	NA
PY 1	$897	$890	$948	$954	$5.6 ($5.4)	(-$3.3, $14.4)
PY 2	$949	$947	$1,008	$1,013	$6.1 ($7.0)	(-$5.4, $17.6)
PY 3	$989	$996	$1,038	$1,051	$3.0 ($6.8)	(-$8.1, $14.1)
PY 4	$946	$960	$1,013	$1,005	$3.4 ($7.6)	(-$9.1, $16.0)
PY 5	$1,034	$1,046	$1,102	$1,084	$10.6 ($8.2)	(-$3.0, $24.1)
PY 1 through PY 5	$965	$972	$1,020	$1,020	$5.7 ($5.6)	(-$3.6, $14.9)
Service use (per 1,000 beneficiaries per year)						
Acute hospitalizations (short-stay acute care and critical access hospitals)						
Baseline	292	288	319	307	NA	NA
PY 1	292	287	320	308	-0.2 (2.4)	(-4.1, 3.7)
PY 2	289	286	322	310	-0.7 (2.9)	(-5.4, 4.0)
PY 3	286	287	314	305	-2.3 (2.8)	(-6.9, 2.3)
PY 4	245	247	275	259	0.5 (2.7)	(-4.1, 5.0)
PY 5	246	246	278	254	-2.3 (2.9)	(-7.0, 2.4)
PY 1 through PY 5	270	269	302	288	-1.0 (2.2)	(-4.7, 2.7)
Outpatient ED visits, including observation stays						
Baseline	492	492	565	552	NA	NA
PY 1	486	490	561	553	-0.8 (4.2)	(-7.7, 6.1)
PY 2	483	489	550	539	-3.3 (4.9)	(-11.4, 4.8)
PY 3	483	488	544	532	-4.8 (5.2)	(-13.3, 3.7)
PY 4	378	384	429	421	6.1 (5.4)	(-2.8, 15.1)
PY 5	408	415	463	449	3.5 (6.0)	(-6.4, 13.4)
PY 1 through PY 5	445	450	511	500	-0.2 (4.3)	(-7.3, 7.0)

Table 5.A.5.4a. ***(continued)***

	Track 2 – Overall					
	CPC+ mean[a]	Comparison mean[a]	Non-CPC+ mean[a]	Non-comparison mean[a]	Impact estimate[b] (SE)	90% confidence interval
Total urgent care center (UCC) visits						
Baseline	97	106	94	94	NA	NA
PY 1	111	18	105	106	1.9 (2.7)	(-2.6, 6.4)
PY 2	124	130	111	110	1.9 (3.9)	(-4.6, 8.4)
PY 3	134	145	122	125	1.3 (5.2)	(-7.2, 9.9)
PY 4	136	140	125	122	3.4 (5.1)	(-4.9, 11.7)
PY 5	186	189	169	162	2.6 (8.1)	(-10.7, 15.9)
PY 1 through PY 5	140	146	126	124	2.2 (3.9)	(-4.3, 8.6)
Ambulatory primary care visits						
Baseline	4,361	4,438	4,597	4,651	NA	NA
PY 1	4,364	4,498	4,666	4,753	-42.6* (22.1)	(-79.0, -6.3)
PY 2	4,393	4,514	4,677	4,753	-33.9 (29.7)	(-82.8, 15.0)
PY 3	4,449	4,557	4,700	4,782	-14.9 (35.9)	(-74.0, 44.1)
PY 4	4,019	4,152	4,480	4,381	-47.5 (52.2)	(-133.4, 38.4)
PY 5	4,236	4,398	4,657	4,595	61.4 (66.4)	(-47.8, 170.7)
PY 1 through PY 5	4,286	4,417	4,637	4,654	-17.0 (29.5)	(-65.4, 31.5)
Proportion of ambulatory primary care visits that were not face-to-face						
PY 4	16.9%	15.6%	15.7%	14.7%	0.6 (0.5)	(-0.2, 1.4)
PY 5	8.9%	8.0%	8.9%	7.9%	0.2 (0.3)	(-0.3, 0.7)
Unweighted sample sizes						
Number of practices	1,515	3,783	7,276	20,115		
Number of beneficiaries	1,896,880	4,507,499	3,378,353	11,153,265		
Number of beneficiary years[c]	7,225,289	17,054,519	12,425,158	42,183,175		

Source: Mathematica's analysis of Medicare claims data from January 2013 through December 2021.

Notes: Although this table indicates which estimates are statistically significant, when we interpret evidence, we combine evidence from the magnitude of the effect, the *p*-values, findings on related outcomes, and sensitivity tests.

[a] We report the actual, unadjusted CPC+ mean for each time period shown in the table. For comparison group practices, non-CPC+ practices, and non-comparison practices, we report the actual, unadjusted mean during the baseline period but the adjusted mean during each intervention period. We obtained the adjusted mean by subtracting the regression-adjusted difference between the CPC+ mean and each group's mean in each time period from the CPC+ mean in that same time period.

Table 5.A.5.4a. ***(continued)***

[b] Impact estimates are regression-adjusted for pre-CPC+ beneficiary characteristics (including HCC scores), practice fixed effects, and COVID-19 controls. Each impact estimate except for the non-face-to-face percentage of ambulatory primary care visits outcome is based on a triple-differences analysis and reflects the difference between (1) the difference in the regression-adjusted average outcome for attributed Medicare FFS beneficiaries in CPC+ practices in the five years of CPC+ compared with baseline relative to the same difference over time for attributed Medicare FFS beneficiaries in comparison practices, and (2) the difference in the regression-adjusted average outcome for attributed Medicare FFS beneficiaries in non-CPC+ practices in the five years of CPC+ compared with baseline relative to the same difference over time for attributed Medicare FFS beneficiaries in non-comparison practices. Impact estimates for the non-face-to-face percentage of ambulatory primary care visits outcome reflect the difference between (1) the difference in the regression-adjusted average outcome for attributed Medicare FFS beneficiaries in CPC+ practices and comparison practices in PY 4 (or PY 5), and (2) the difference in the regression-adjusted average outcome for attributed Medicare FFS beneficiaries in non-CPC+ practices and non-comparison practices in PY 4 (or PY 5).

[c] After accounting for weights that adjust for matching, time observed in Medicare FFS, and the concentration of CPC+ in each geographic area, the effective sample sizes are reduced. For non-CPC+ practices, the effective sample size (in terms of beneficiary-years) is 28 percent of the actual group size. For non-comparison practices, the effective sample size (in terms of beneficiary-years) is 12 percent of the actual group size. For the comparison group, the effective sample size (in terms of beneficiary-years) is 40 percent of the size of the actual comparison group. Because CPC+ sample size is affected only by time the beneficiary is observed (and is not affected by the matching weights), the effective sample size for the CPC+ group is about 96 percent of the actual sample size.

*/**/*** Significantly different from zero at the 0.10/0.05/0.01 level, two-tailed test.

ED = emergency department; FFS = fee-for-service; HCC = hierarchical condition category; NA = not applicable; non-comparison = unselected practices in CPC+ comparison regions; non-CPC+ = non-participating practices in CPC+ regions; PY = Program Year; SE = standard error.

Table 5.A.5.4b. Regression-adjusted means and estimated triple-differences and difference-in-differences impacts of CPC+ on selected expenditures and service use outcomes for attributed Medicare FFS beneficiaries by program year and average across the five program years, Track 2 by SSP status

	Track 2 – SSP						Track 2 – Non-SSP					
	CPC+ mean[a]	Comparison mean[a]	Non-CPC+ mean[a]	Non-comparison mean[a]	Impact estimate[b] (SE)	90% confidence interval	CPC+ mean[a]	Comparison mean[a]	Non-CPC+ mean[a]	Non-comparison mean[a]	Impact estimate[b] (SE)	90% confidence interval
Medicare expenditures (per beneficiary per month)												
Total Medicare Part A and B expenditures without enhanced payments for CPC+ and SSP												
Baseline	896	893	970	956	NA	NA	861	865	895	906	NA	NA
PY 1	917	910	990	982	$4.7 ($8.9)	(-$10.0, $19.3)	881	875	915	925	$5.4 ($6.2)	(-$4.9, $15.7)
PY 2	966	969	1,049	1,041	-$0.3 ($12.2)	(-$20.5, $19.8)	935	929	975	985	$10.1 ($7.5)	(-$2.2, $22.4)
PY 3	1,009	1,020	1,068	1,077	$3.2 ($11.3)	(-$15.4, $21.8)	974	977	1,014	1,025	$1.2 ($7.8)	(-$11.7, $14.0)
PY 4	956	983	1,065	1,026	-$2.9 ($12.5)	(-$23.5, $17.7)	938	942	970	986	$9.7 ($8.4)	(-$4.1, $23.6)
PY 5	1,048	1,073	1,176	1,102	-$8.4 ($14.2)	(-$31.8, $15.0)	1,022	1,025	1,036	1,065	$21.8** ($9.0)	($7.0, $36.5)
PY 1 through PY 5	982	995	1,069	1,044	-$0.3 ($9.4)	(-$15.8, $15.2)	952	954	980	996	$9.1 ($6.3)	(-$1.2, $19.4)
Service use (per 1,000 beneficiaries per year)												
Acute hospitalizations (short-stay acute care and critical access hospitals)												
Baseline	300	291	339	309	NA	NA	287	286	304	305	NA	NA
PY 1	302	291	339	311	0.1 (3.9)	(-6.2, 6.5)	285	284	304	305	-0.7 (2.8)	(-5.4, 4.0)
PY 2	297	289	342	312	-0.5 (4.8)	(-8.5, 7.4)	282	284	306	308	-1.1 (3.2)	(-6.4, 4.2)
PY 3	296	291	331	308	1.5 (4.6)	(-6.0, 9.0)	278	284	301	302	-6.0* (3.3)	(-11.4, -0.6)
PY 4	253	252	300	263	3.0 (4.5)	(-4.4, 10.3)	239	244	254	256	-2.0 (3.2)	(-7.3, 3.3)
PY 5	256	252	312	255	-5.8 (4.8)	(-13.7, 2.1)	237	241	249	253	-0.6 (3.2)	(-5.9, 4.7)
PY 1 through PY 5	280	274	325	290	-0.2 (3.7)	(-6.3, 5.8)	263	266	284	285	-2.1 (2.6)	(-6.4, 2.2)

Table 5.A.5.4b. ***(continued)***

	Track 2 – SSP						Track 2 – Non-SSP					
	CPC+ mean[a]	Comparison mean[a]	Non-CPC+ mean[a]	Non-comparison mean[a]	Impact estimate[b] (SE)	90% confidence interval	CPC+ mean[a]	Comparison mean[a]	Non-CPC+ mean[a]	Non-comparison mean[a]	Impact estimate[b] (SE)	90% confidence interval
Outpatient ED visits, including observation stays												
Baseline	479	475	545	542	NA	NA	502	506	581	563	NA	NA
PY 1	471	472	541	544	-1.1 (7.2)	(-13.0, 10.8)	498	504	576	562	-1.0 (4.6)	(-8.6, 6.6)
PY 2	467	472	531	528	-7.2 (8.6)	(-21.3, 7.0)	496	502	565	550	-0.5 (5.3)	(-9.2, 8.2)
PY 3	468	472	523	518	-10.5 (8.8)	(-24.9, 4.0)	495	501	560	547	0.1 (5.8)	(-9.4, 9.7)
PY 4	362	375	421	412	-3.8 (8.5)	(-17.8, 10.3)	391	391	436	430	15.4** (6.5)	(4.6, 26.1)
PY 5	392	408	465	440	-2.4 (10.0)	(-18.9, 14.1)	420	421	461	459	10.7 (6.8)	(-0.5, 21.9)
PY 1 through PY 5	430	437	497	489	-5.1 (7.2)	(-17.0, 6.8)	457	461	522	510	4.3 (4.9)	(-3.8, 12.3)
Total urgent care center (UCC) visits												
Baseline	99	103	89	98	NA	NA	96	107	99	90	NA	NA
PY 1	115	117	104	111	1.0 (5.0)	(-7.2, 9.1)	108	119	106	101	2.2 (2.8)	(-2.4, 6.9)
PY 2	132	129	116	114	-1.5 (6.7)	(-12.5, 9.6)	118	131	107	106	4.1 (4.4)	(-3.1, 11.3)
PY 3	138	146	128	132	-3.8 (9.5)	(-19.4, 11.8)	131	145	118	119	4.5 (5.0)	(-3.7, 12.7)
PY 4	141	142	128	147	-2.1 (8.3)	(-15.7, 11.5)	131	138	122	116	8.0 (5.6)	(-1.2, 17.2)
PY 5	196	198	171	173	-0.1 (15.7)	(-25.9, 25.8)	179	181	167	150	2.3 (8.7)	(-12.0, 16.7)
PY 1 through PY 5	146	148	129	131	-1.3 (6.9)	(-12.6, 10.0)	135	145	123	118	4.2 (4.3)	(-3.0, 11.3)

Table 5.A.5.4b. ***(continued)***

	Track 2 – SSP						Track 2 – Non-SSP					
	CPC+ mean[a]	Comparison mean[a]	Non-CPC+ mean[a]	Non-comparison mean[a]	Impact estimate[b] (SE)	90% confidence interval	CPC+ mean[a]	Comparison mean[a]	Non-CPC+ mean[a]	Non-comparison mean[a]	Impact estimate[b] (SE)	90% confidence interval
Ambulatory primary care visits												
Baseline	4,214	4,355	4,494	4,526	NA	NA	4,476	4,504	4,678	4,777	NA	NA
PY 1	4,237	4,416	4,536	4,638	7.2 (32.5)	(-46.3, 60.7)	4,466	4,564	4,769	4,869	-83.8*** (29.6)	(-132.4, -35.1)
PY 2	4,268	4,441	4,578	4,645	-30.1 (45.5)	(-104.9, 44.8)	4,494	4,572	4,755	4,960	-37.6 (38.4)	(-100.8, 25.7)
PY 3	4,333	4,497	4,610	4,681	-18.9 (56.9)	(-112.5, 74.6)	4,542	4,605	4,771	4,882	-14.3 (44.8)	(-88.1, 59.4)
PY 4	3,913	4,115	4,641	4,297	-70.6 (82.2)	(-205.8, 64.7)	4,103	4,181	4,346	4,464	-16.2 (49.2)	(-97.2, 64.7)
PY 5	4,161	4,375	4,835	4,523	112.9 (100.8)	(-53.0, 278.8)	4,295	4,417	4,501	4,668	-26.0 (54.7)	(-116.0, 64.0)
PY 1 through PY 5	4,178	4,363	4,639	4,558	-1.7 (44.3)	(-74.6, 71.3)	4,373	4,459	4,635	4,750	-36.9 (36.9)	(-97.7, 23.9)
Proportion of ambulatory primary care visits that were not face-to-face												
PY 4	0%	0%	0%	0%	-0.4 (0.8)	(-1.7, 0.9)	0%	0%	0%	0%	0.8 (0.6)	(-0.1, 1.8)
PY 5	0%	0%	0%	0%	-0.1 (0.5)	(-0.8, 0.7)	0%	0%	0%	0%	0.4 (0.4)	(-0.2, 1.0)
Unweighted sample sizes												
Number of practices	636	1,817	2,423	5,010			879	1,966	4,853	15,105		
Number of beneficiaries	847,208	2,257,322	1,375,874	3,618,373			1,053,634	2,261,852	2,026,657	7,636,397		
Number of beneficiary years[c]	3,204,963	8,538,135	4,985,186	13,270,465			4,020,326	8,516,384	7,439,972	28,912,710		

Source: Mathematica's analysis of Medicare claims data from January 2013 through December 2021.

Notes: Although this table indicates which estimates are statistically significant, when we interpret evidence, we combine evidence from the magnitude of the effect, the *p*-values, findings on related outcomes, and sensitivity tests.

[a] We report the actual, unadjusted CPC+ mean for each time period shown in the table. For comparison group practices, non-CPC+ practices, and non-comparison practices, we report the actual, unadjusted mean during the baseline period but the adjusted mean during each intervention period. We obtained the adjusted mean by subtracting the regression-adjusted difference between the CPC+ mean and each group's mean in each time period from the CPC+ mean in that same time period.

Table 5.A.5.4b. ***(continued)***

[b] Impact estimates are regression-adjusted for pre-CPC+ beneficiary characteristics (including HCC scores), practice fixed effects, and COVID-19 controls. Each impact estimate except for the non-face-to-face percentage of ambulatory primary care visits outcome is based on a triple-differences analysis and reflects the difference between (1) the difference in the regression-adjusted average outcome for attributed Medicare FFS beneficiaries in CPC+ practices in the five years of CPC+ compared with baseline relative to the same difference over time for attributed Medicare FFS beneficiaries in comparison practices, and (2) the difference in the regression-adjusted average outcome for attributed Medicare FFS beneficiaries in non-CPC+ practices in the five years of CPC+ compared with baseline relative to the same difference over time for attributed Medicare FFS beneficiaries in non-comparison practices. Impact estimates for the non-face-to-face percentage of ambulatory primary care visits outcome reflect the difference between (1) the difference in the regression-adjusted average outcome for attributed Medicare FFS beneficiaries in CPC+ practices and comparison practices in PY 4 (or PY 5), and (2) the difference in the regression-adjusted average outcome for attributed Medicare FFS beneficiaries in non-CPC+ practices and non-comparison practices in PY 4 (or PY 5).

[c] After accounting for weights that adjust for matching, time observed in Medicare FFS, and the concentration of CPC+ in each geographic area, the effective sample sizes are reduced. For the SSP group: for non-CPC+ practices, the effective sample size (in terms of beneficiary-years) is 22 percent of the actual group size. For non-comparison practices, the effective sample size (in terms of beneficiary-years) is 12 percent of the actual group size. For the comparison group, the effective sample size (in terms of beneficiary-years) is 38 percent of the size of the actual comparison group. Because CPC+ sample size is affected only by time the beneficiary is observed (and is not affected by the matching weights), the effective sample size for the CPC+ group is about 96 percent of the actual sample size. For the non-SSP group: for non-CPC+ practices, the effective sample size (in terms of beneficiary-years) is 39 percent of the actual group size. For non-comparison practices, the effective sample size (in terms of beneficiary-years) is 18 percent of the actual group size. For the comparison group, the effective sample size (in terms of beneficiary-years) is 43 percent of the size of the actual comparison group. Because CPC+ sample size is affected only by time the beneficiary is observed (and is not affected by the matching weights), the effective sample size for the CPC+ group is about 96 percent of the actual sample size.

*/**/*** Significantly different from zero at the 0.10/0.05/0.01 level, two-tailed test.

ED = emergency department; FFS = fee-for-service; HCC = hierarchical condition category; n.a. = not applicable; non-comparison = unselected practices in CPC+ comparison regions; non-CPC+ = non-participating practices in CPC+ regions; PY = Program Year; SE = standard error; SSP = Medicare Shared Savings Program.

Table 5.A.5.5. Estimated average annual impacts of CPC+ across the five program years on Medicare expenditures without enhanced payments, from triple-differences main analysis and sensitivity tests, Tracks 1 and 2

	Track 1				Track 2			
	Average annual impact estimate[a] (SE)	Average annual percentage impact[b]	90% confidence interval	*p*-Value	Average annual impact estimate[a] (SE)	Average annual percentage impact[b]	90% confidence interval	*p*-Value
Main triple-differences estimates								
PY 1 through PY 5	6.47 (4.96)	0.7%	(-1.68, 14.62)	0.19	5.66 (5.61)	0.6%	(-3.57, 14.89)	0.31
Excluding COVID-19 controls								
PY 1 through PY 5	5.48 (5.01)	0.6%	(-2.77, 13.72)	0.28	5.23 (5.65)	0.5%	(-4.07, 14.53)	0.36
With winsorized concentration weights at the 99th percentile								
PY 1 through PY 5	8.67 (4.66)	0.9%	(1.01, 16.33)	0.06	8.36 (5.10)	0.9%	(-0.04, 16.75)	0.10
Without concentration weight for non-CPC+ practices and non-comparison practices								
PY 1 through PY 5	9.91 (3.81)	1.0%	(3.64, 16.18)	0.01	9.97 (4.34)	1.0%	(2.83, 17.10)	0.02
Excluding practices that share the same TIN as CPC+ or comparison practices								
PY 1 through PY 5	12.90 (5.45)	1.3%	(3.94, 21.86)	0.02	10.85 (6.23)	1.1%	(0.60, 21.10)	0.08
Include only beneficiaries attributed in first quarter of baseline and intervention periods								
PY 1 through PY 5	8.14 (5.20)	0.8%	(-0.41, 16.70)	0.12	7.76 (5.63)	0.8%	(-1.50, 17.02)	0.17

Source: Mathematica's analysis of Medicare claims data from January 2013 through December 2021.

Notes: Although this table indicates which estimates are statistically significant, when we interpret evidence, we combine evidence from the magnitude of the effect, the *p*-values, findings on related outcomes, and sensitivity tests.

[a] Impact estimates are regression-adjusted for pre-CPC+ beneficiary characteristics (including HCC scores), practice fixed effects, and COVID-19 controls. Each impact estimate is based on a triple-differences analysis and reflects the difference between (1) the difference in the regression-adjusted average outcome for attributed Medicare FFS beneficiaries in CPC+ practices in the five years of CPC+ compared with baseline relative to the same difference over time for attributed Medicare FFS beneficiaries in comparison practices, and (2) the difference in the regression-adjusted average outcome for attributed Medicare FFS beneficiaries in non-CPC+ practices in the five years compared with baseline relative to the same difference over time for attributed Medicare FFS beneficiaries in non-comparison practices.

[b] We calculate percentage impacts relative to what the CPC+ mean would have been in each year in the absence of the intervention—that is, the unadjusted CPC+ mean minus the impact estimate.

*/**/*** Significantly different from zero at the 0.10/0.05/0.01 level, two-tailed test.

FFS = fee-for-service; HCC = hierarchical condition category; non-comparison = unselected practices in CPC+ comparison regions; non-CPC+ = non-participating practices in CPC+ regions; PY = Program Year; SE = standard error; TIN = Tax Identification Number.

Table 5.A.5.6. *P*-values for the joint significance test of the difference between difference-in-differences and triple-differences estimated impacts of CPC+ on Medicare expenditures without enhanced payments, acute hospitalizations, and outpatient ED visits, by program year and average across the five program years, Tracks 1 and 2

	Track 1			Track 2		
Program year	Overall	SSP	Non-SSP	Overall	SSP	Non-SSP
PY 1	0.38	0.64	0.19	0.21	0.62	0.24
PY 2	0.36	0.32	0.22	0.88	0.97	0.38
PY 3	0.41	0.35	0.02	0.67	0.50	0.19
PY 4	0.00	0.05	0.05	0.00	0.04	0.01
PY 5	0.18	0.42	0.18	0.00	0.02	0.09
PY 1 through PY 5	0.02	0.20	0.04	0.00	0.00	0.13

Source: Mathematica's analysis of Medicare claims data from January 2013 through December 2021.

Notes: We report the joint statistical significance of the difference between the impact estimates from the main impact analysis and the triple-differences analysis, across the three priority outcomes (Medicare expenditures excluding fees, acute hospitalizations, and outpatient ED visits), within and across program years using a Seemingly Unrelated Regressions model.

ED = emergency department; PY = Program Year; SSP = Medicare Shared Savings Program.

5.B. Attribution methodology

In this Appendix, we explain beneficiary attribution (Section 1), describe each step of the attribution approach we use for CPC+ and comparison practices (Section 2), and discuss how the methodology has changed over time (Section 3). We then compare how our evaluation attribution process differs from CMS's payment attribution (Section 4). Finally, we explore similarities between our evaluation attribution sample and CMS's payment attribution sample (Section 5). We updated the reported number of attributed beneficiaries, by quarter or year, based on the results from attribution for the final report.

5.B.1. What is beneficiary attribution?

Attribution is a methodology used to identify the population of beneficiaries under the care of a particular practitioner, practice, or health system. CPC+ provides each participating practice site with enhanced and alternative payments for their Medicare fee-for-service (FFS) beneficiaries. A practice site is composed of a unique grouping of practitioners and billing numbers (described in more detail below). To determine the amount of payments that practices receive, CMS uses attribution to measure the number and acuity of the Medicare FFS population receiving regular, continuous care from the practice. The CPC+ payment attribution process uses Medicare administrative data (claims and enrollment data) to identify the Medicare FFS beneficiaries associated with CPC+ practices.[2,3]

As a part of the evaluation of CPC+, we use a similar claims-based attribution process to assign Medicare beneficiaries to all primary care practice sites serving Medicare beneficiaries in a given quarter. We run our own attribution so we can attribute Medicare beneficiaries to both CPC+ and comparison practices using an identical methodology. We assign eligible Medicare beneficiaries to practice sites for each quarter of the time period we are analyzing. For the final report, this period includes 4 baseline quarters in 2016 and 20 intervention quarters in 2017, 2018, 2019, 2020, and 2021 for the 2017 Starters.[4] Although we use a process similar to CMS payment attribution, there are a few key differences that we highlight in Section 5.B.4.

5.B.2. How do we do attribution?

Like the CMS payment attribution method, attribution for the CPC+ evaluation uses Medicare administrative data to assign Medicare FFS beneficiaries to CPC+ and comparison practice sites. The CPC+ evaluation attribution process consists of five steps. First, we identify a pool of primary care practices that compete for beneficiaries in the attribution process. Second, because we use Medicare claims, which report the practitioners who provided the service rather than the practice, we group practitioners into the practices identified in the first step. Third, we identify the set of beneficiaries who are eligible for attribution. Fourth, we identify the set of primary care services that we consider in the

[2] See CMS's CPC+ Payment Methodologies at https://innovation.cms.gov/media/document/cpc-plus-payment-methodology-cy2021 for details on CPC+ payment attribution (Chapter 2). In Section 5.B.4 below, we summarize the differences between the payment and evaluation attribution processes.

[3] Starting in 2019, CMS incorporated Voluntary Alignment, a method by which beneficiaries confirm their primary care practitioner, into CPC+ attribution methodology.

[4] Beneficiaries are assigned to the first practice they are attributed to in that period (i.e., the baseline or the intervention period).

attribution process. Fifth, we use the information from the previous four steps to attribute eligible Medicare beneficiaries to a single practice in each quarter.

Below we describe each of these steps in detail.

Step 1: Identify a pool of primary care practices

To develop a frame of primary care practices that compete for beneficiaries in the attribution process, we start with a roster of all practices in the United States with at least one practitioner (defined as a physician, nurse practitioner, or physician assistant) with a primary care specialty (defined as family practice, general practice, geriatrics, or internal medicine). We purchase yearly rosters from IQVIA, a commercial health care data vendor that maintains and verifies lists of practitioners who work in practices throughout the country, including practices' names and addresses along with the name, specialty, and National Provider Identifier (NPI) of each practitioner at the practice site.[5] We augment the IQVIA data with practitioner taxonomy and Medicare specialty codes and fill in missing NPIs by linking the practitioner-level IQVIA data to the National Plan and Provider Enumeration System (NPPES). We then identify CPC+ practices within the roster of IQVIA practices, using a combination of address, name, and practitioner matching. If we cannot identify a CPC+ practice in the IQVIA roster, we augment the IQVIA data by appending CPC+ practice and practitioner data from CMS.

Step 2: Group practitioners into practice sites

Two key inputs in attribution are a roster of practitioners working at practice sites and the information they use to bill Medicare for services provided at those practice sites. In the CMS payment attribution method for CPC+, a practice is defined by the combinations of Taxpayer Identification Number (TIN) (or CMS Certification Number [CCN] for critical access hospitals) and NPIs identified for each practitioner at the practice site. Participating CPC+ practices submit this information in monthly rosters. Each service in the Medicare claims data includes (1) the TIN or CCN and (2) the NPI of the practitioner who rendered the service. CMS determines whether the TIN (or CCN) and NPI combination on the claim match a TIN (or CCN) and NPI combination in a practitioner-practice site roster. If so, the visit is associated with that practice in the CPC+ payment attribution algorithm. Otherwise, CMS assigns that visit to the individual practitioner identified as the single TIN-NPI or CCN-NPI combination.

To facilitate attribution for the evaluation, we proceed with three substeps to construct a roster of practitioners working at all CPC+ and potential comparison practices and their associated TINs (or CCNs) and NPIs.

Substep 1: Create initial roster of NPIs from yearly rosters

As a starting point, we use practitioner rosters we purchased from IQVIA for years 2016 through 2021, which provide the practices' roster of practitioners in that year (we use the 2016 roster for the period 2014 through 2016).[6] The rosters connect a unique practice ID to a list of practitioners in each year. Although we had extensive information about CPC+ practices from their applications, for matching purposes, we opted to identify CPC+ practice and practitioner characteristics using the same data source (IQVIA) as we

[5] The purchased yearly rosters were based on SK&A data for the baseline period, PY 1, and PY 2 of CPC+. Starting in 2019, IQVIA discontinued the SK&A data and replaced it with OneKey data. For PY 3 through PY 5, the purchased yearly rosters are based on the OneKey database.

[6] Our attribution process uses a two-year lookback period, so we need practitioner rosters for 2014 onward.

used for the potential comparison practices, both at baseline and over time. This approach removes bias that could result from using different data sources for the two groups, such as more frequent or thorough updates to practitioner rosters in the CPC+ data than in IQVIA data. Over the six-year period examined in the final report, we found that the IQVIA roster captured 74.3 to 85.3 percent of practitioners in the CPC+ rosters. This finding suggests that, although IQVIA data are not perfectly capturing CPC+ practitioners, our rosters include a high proportion of them. We explore this topic more extensively in Section 5.B.5.

Substep 2: Assign TINs to each practice in roster

Because the IQVIA data do not include the practice or practitioner TINs used in the payment attribution method, we use claims data to assign TINs to each practice.[7] To do so, we use an algorithm that picks the TIN most frequently billed in Medicare claims data for primary care services by the NPIs of primary care practitioners that the IQVIA roster indicates are located at a practice.[8] We start by assigning a single TIN to a practice in each year over the seven-year period from 2015 through 2021.[9] We then maintain all TINs previously associated with a practice, resulting in practices with multiple TINs at a given time. Additionally, we backdate the start date of each TIN by one calendar year to ensure we correctly associate claims billed by a practice at some point during the year prior to the practice's new TIN.[10]

Substep 3: Unique NPI/TIN assignment

In some instances, the same NPI and TIN combination occurs at multiple practices identified in the IQVIA data at the same time (approximately 18 percent of all practice-practitioner observations share the same NPI and TIN in the 2021 roster). This occurs when a practitioner works in more than one practice site within a health care system (if the practice sites share the same billing TIN [including historic TINs]). In these cases, we cannot distinguish which practice provided care for a beneficiary. To reconcile duplicate NPI–TIN combinations before attribution, we assign the NPI to one practice using the following hierarchy of rules: (1) if the duplicate occurs between a CPC+ practice and a comparison practice, we assign the duplicate to the CPC+ practice; (2) ascending practice size, as measured by number of primary care practitioners (that is, we assign

[7] When developing our method of assigning TINs to practices, we used CPC+ application data to assess the accuracy of the approach. For CPC+ applicants, we examined the overlap between the assigned TINs and reported TINs: for 95 percent of applicants, at least one assigned TIN was also on the CPC+ application. Using the assigned TINs in attributing beneficiaries to CPC+ practices (rather than using TINs on the CPC+ application) increases the risk of misattributing beneficiaries to CPC+ practices (if we assigned an incorrect or invalid TIN to that practice).

[8] In practices where at least one practitioner is found to practice only at that practice per the IQVIA data, we limit practitioners used in TIN assignment to these "single-site" practitioners. For practices where there are no single-site practitioners, we use all primary care practitioners associated with the practice in TIN assignment.

[9] We decided not to do TIN assignment for 2014, because we would have had to use a very out-of-date roster (one from October 2016). We were concerned that this would cause a misspecification of the TIN. Since we maintain all TINs previously associated with the practice, we did not want to include a potentially misspecified TIN that would be included in all subsequent years. Note, however, that we backdate the TIN assigned in 2015 to 2014.

[10] Specifically, we backdate assigned TINs in this way to avoid cases where the practice switched ownership (and so the TIN changed) midyear. Because we use a plurality approach to assigning TINs to a year, if we did not backdate TINs (for example, by forcing only one TIN to be active during a year) we would not assign the correct practice on up to 50 percent of the claims for that switching year.

the NPI to the smaller practice); and (3) random assignment, if the duplicate occurs among practices in the same research group (CPC+ or potential comparison) and of the same size.[11]

This process results in a master practitioner file with a unique crosswalk between NPIs-TINs and their associated practice IDs in each year. We use this crosswalk to map each Medicare service to a particular practice.

Step 3: Identify Medicare beneficiaries eligible for attribution

We start with the list of beneficiaries who had at least one primary care visit (see Step 4 for definition of primary care visits) to any NPI in our master practitioner file (created in Step 2). We then limit the pool of beneficiaries to those who meet the eligibility criteria. To be eligible for evaluation attribution in a given quarter, beneficiaries must meet the following criteria at the start of the quarter, as indicated by the Medicare enrollment database (EDB):[12,13]

1. Be enrolled in both Medicare Part A and Part B,
2. Have Medicare as their primary payer,
3. Not be covered under a Medicare Advantage or other Medicare health plan,
4. Not be incarcerated,
5. Be alive.

These criteria ensure that we can reliably measure beneficiary outcomes in the Medicare FFS data unlike, for example, beneficiaries enrolled in a Medicare Advantage plan.

Step 4: Identify primary care claims used in attribution

We next narrow the universe of all billed Medicare services to the primary care services used in beneficiary attribution. There are four criteria for a billed service that determine whether we use it in attribution for a given quarter: (1) type of claim, (2) date of the claim, (3) type of service, and (4) practitioner. A service must meet all four criteria to be included in the attribution process.

1. Type of claim

For attribution, we use national Medicare FFS Physician and Outpatient claims. Most visits are in the Physician file, except claims submitted by critical access hospitals, which are in the Outpatient file.

[11] Consistent with CMS's attribution approach, we prioritize the smaller practice to avoid dropping any practices altogether.

[12] For example, beneficiaries must meet all eligibility criteria on January 1, 2017, to be eligible for evaluation attribution in the first quarter of 2017 (January 1, 2017–March 31, 2017).

[13] The EDB provides information, by month, for beneficiaries enrolled in Medicare, including the parts of Medicare in which they were enrolled—Part A, Part B, or Part C (a health maintenance organization), whether Medicare was their primary payer of medical bills, whether they were incarcerated, and the date they died, if applicable.

2. Date of the claim

We use primary care services that occurred during a 24-month "lookback" period in the attribution process. For each quarter, the lookback period is the 24-month period that ended immediately before the quarter started. For example, we use claims from January 2015 to December 2016 to attribute beneficiaries to CPC+ practices for the first quarter of 2017. Table 5.B.1 lists the lookback periods we used for each quarter in the annual report. Claims for attribution were pulled on May 3, 2018, for the first through fourth quarters of 2016, on March 20, 2020, for the first quarter of 2017 through the fourth quarter of 2018, and on March 11, 2022, for the first quarter of 2019 through the fourth quarter of 2021.

Table 5.B.1. Lookback periods for annual report quarterly beneficiary attribution

Attribution quarter	CPC+ period for 2017 Starters	Lookback period
2016 Q1	Baseline	Jan. 2014–Dec. 2015
2016 Q2	Baseline	Apr. 2014–Mar. 2016
2016 Q3	Baseline	July 2014–June 2016
2016 Q4	Baseline	Oct. 2014–Sept. 2016
2017 Q1	Intervention	Jan. 2015–Dec. 2016
2017 Q2	Intervention	Apr. 2015–Mar. 2017
2017 Q3	Intervention	July 2015–June 2017
2017 Q4	Intervention	Oct. 2015–Sept. 2017
2018 Q1	Intervention	Jan. 2016–Dec. 2017
2018 Q2	Intervention	Apr. 2016–Mar. 2018
2018 Q3	Intervention	July 2016–June 2018
2018 Q4	Intervention	Oct. 2016–Sept. 2018
2019 Q1	Intervention	Jan. 2017–Dec. 2018
2019 Q2	Intervention	Apr. 2017–Mar. 2019
2019 Q3	Intervention	July 2017–June 2019
2019 Q4	Intervention	Oct. 2017–Sept. 2019
2020 Q1	Intervention	Jan. 2018–Dec. 2019
2020 Q2	Intervention	Apr. 2018–Mar. 2020
2020 Q3	Intervention	July 2018–June 2020
2020 Q4	Intervention	Oct. 2018–Sept. 2020
2021 Q1	Intervention	Jan. 2019–Dec. 2020
2021 Q2	Intervention	Apr. 2019–Mar. 2021
2021 Q3	Intervention	July 2019–June 2021
2021 Q4	Intervention	Oct. 2019–Sept. 2021

Q = quarter.

3. Type of service

Next, we limit claims to eligible primary care services using the Current Procedural Terminology (CPT) code reported on the claim. Table 5.B.2 lists the CPT codes of services that we consider to be related to primary care, following the definition CMS uses for CPC+ payment attribution. A subset of eligible primary care services are related to chronic care management (CCM); these claims receive precedence in the attribution algorithm (described below). For the 2020 and 2021 quarters, we examined the potential effects of coronavirus 2019 (COVID-19) on evaluation attribution, and how including telehealth procedure codes in the attribution algorithm might alter those effects. We found that using telehealth codes for attribution led to a very small increase in the number of attributed beneficiaries (close to 0 percent in the second quarter of 2020 and up to 0.5 percent in the last quarter of 2021, in both CPC+ and comparison practices). Therefore, we decided not to include telehealth codes in the evaluation attribution, which is consistent with CMS's decision for payment attribution for 2020 and 2021 quarters (and past quarters as well).

Table 5.B.2. Primary care services eligible for attribution

Type of service	Service	CPT codes
All primary care	Office/outpatient visit evaluation and management (E&M)	99201–99205 99211–99215
	Home care	99324-99328 99334–99337 99339–99345 99347–99350
	Welcome to Medicare and Annual Wellness visits	G0402, G0438, G0439
	Advance care planning	99497
	Collaborative care model	G0502–G0504[a] 99492, 99493, 99494[b]
	Cognition and functional assessment for patient with cognitive impairment	G0505[a], 99483[b]
	Outpatient clinic visit for assessment and management (CAHs only)	G0463
	Transitional care management services	99495–99496
CCM-related service	CCM services	99490, 99491[c]
	Complex CCM services	99487, 99488[d]
	Assessment/care planning for patients requiring CCM services	G0506[a]
	Care management services for behavioral health conditions	G0507[a], 99484[b]
	Prolonged services without face-to-face contact	99358[a]

[a] Added effective January 1, 2017.

[b] Added effective January 1, 2018.

[c] Added effective January 1, 2019.

[d] Discontinued effective January 1, 2017.

CAH = critical access hospital; CCM = chronic care management; CPT = Current Procedural Terminology.

4. Practitioner

Only claims that have a practitioner who is one of the following are included in the attribution process:

- A practitioner in IQVIA data who is part of a practice with at least one practitioner with a primary care specialty (see Steps 1 and 2 for more details).
- A practitioner who is not in IQVIA data but has a primary or secondary primary care specialty determined by the National Plan and Provider Enumeration System (NPPES; see Table 5.B.3 for the list of primary care specialty codes that we and CMS use).
- Any practitioner if the claim is for a CCM service (lower half of Table 5.B.2).

Additionally, we limit claims to services that are reported in the physician (carrier) claims or are from critical access hospitals in the outpatient claims. Like CMS's payment attribution approach, this process excludes claims from federally qualified health centers (FQHCs) and rural health clinics (RHCs).[14]

[14] This restriction means that—in both payment and evaluation attribution—even if beneficiaries have most of their visits at an FQHC or RHC, they would not be attributed to a practice that is an FQHC or RHC.

Table 5.B.3. Primary care practitioner specialties

Primary care specialty	Taxonomy code
Family Medicine	**207Q00000X**
Adult Medicine	207QA0505X
Geriatric Medicine	207QG0300X
Hospice and Palliative Medicine	207QH0002X
General Practice	**208D00000X**
Internal Medicine	**207R00000X**
Geriatric Medicine	207RG0300X
Hospice and Palliative Medicine	207RH0002X
Clinical Nurse Specialist	**364S00000X**
Acute Care	364SA2100X
Adult Health	364SA2200X
Chronic Care	364SC2300X
Community Health/Public Health	364SC1501X
Family Health	364SF0001X
Gerontology	364SG0600X
Holistic	364SH1100X
Women's Health	364SW0102X
Nurse Practitioner	**363L00000X**
Acute Care	363LA2100X
Adult Health	363LA2200X
Community Health	363LC1500X
Family	363LF0000X
Gerontology	363LG0600X
Primary Care	363LP2300X
Women's Health	363LW0102X
Physician Assistant	**363A00000X**
Medical	363AM0700X

Source: CMS's CPC+ Payment Methodologies, at https://innovation.cms.gov/media/document/cpc-plus-payment-methodology-cy2021.

Notes: Blue shading indicates a specialty category. The non-shaded rows are sub-specialties of the prior blue-shaded category.

Step 5: The attribution algorithm

After we identify beneficiaries eligible for attribution and pull all eligible primary care services (as determined by type of claim, date of the claim, type of service, and practitioner), we apply the CPC+ payment attribution algorithm used by CMS. There are three parts to the attribution algorithm:

1. Attribution based on CCM-related billing

If a beneficiary's ***most recent*** eligible primary care visit in the 24-month lookback period was for CCM-related services, we attribute the beneficiary to the practice that provided that CCM-related service.[15]

2. Attribution based on Annual Wellness Visits or Welcome to Medicare visits

Starting in the first quarter of 2018, if a beneficiary is not attributed on the basis of CCM-related billing, and the beneficiary had an Annual Wellness Visit or a Welcome to Medicare visit in the 24-month lookback period, we attribute the beneficiary to the practice that provided the most recent Annual Wellness Visit or a Welcome to Medicare visit.[16]

3. Attribution based on plurality of eligible primary care services

If a beneficiary is not attributed on the basis of Annual Wellness Visits, Welcome to Medicare Visits, or CCM-related billing (including cases in which a beneficiary had CCM billed, but the most recent visit was not for CCM-related services), we count the number of eligible primary care visits the beneficiary received from each practice that provided such services. We then attribute the beneficiary to the practice that provided the plurality (that is, the largest share) of eligible primary care visits during the lookback period. If a beneficiary has the same number of eligible primary care visits at more than one practice, we attribute the beneficiary to the practice where the beneficiary had the most recent visit. If two or more of these practices share the same most recent visit date, we attribute the beneficiary to a practice that is on our IQVIA practitioner roster over a primary care NPI that is not on the roster.[17] We break any further ties randomly.

5.B.3. Changes in attribution methodology across annual reports and across quarters

1. We update data and rerun attribution for quarters in the previous annual report that had updates to the input data (for example, we did this for the 2019 and 2020 quarters in the final report). Other than the data changes, the attribution methodology stays the same between reports for a given quarter.

   Data changes from the fourth to the final report include:

   – Backdating TINs from the 2021 TIN assignment to 2020. This impacted 2020 Quarters 2 through 4, for which we used 2020 claims in the lookback period.

[15] Because CPC+ care management (indicated by the care management fee) and the CCM are duplicative services, it is important to note that CPC+ practices cannot bill for CCM-related services for their CPC+ payment-attributed beneficiaries. CPC+ practices are free to bill for CCM-related services for non-payment-attributed beneficiaries, which may result in future attribution to the CPC+ practice.

[16] We include the Annual Wellness Visit and Welcome to Medicare visit attribution criteria to the attribution algorithm for the first quarter of 2018 onward, to align with the same change CMS made to the CPC+ payment attribution algorithm.

[17] Although, in a tie, CMS payment attribution gives preference to CPC+ practices, we did not want to favor CPC+ practices over comparison practices.

- Updating the 2020 TIN assignment and backdating TINs from the 2020 TIN assignment to 2019. This impacted 2019 Quarter 2 through 2020 Quarter 4, for which we used 2019 and 2020 claims in the lookback period.
- Additional runout of claims, which affected attribution for all quarters in 2019 and 2020.

These data changes resulted in 2019 and 2020 quarters showing slightly different attribution samples in going from the fourth to the final reports.[18]

2. We alter the attribution approach by quarter to reflect relevant changes in CMS's attribution approach, for example, adding the Annual Wellness Visit criteria starting in the first quarter of 2018.

In addition, annual updates to the Health Care Common Procedure Coding System (HCPCS) or other codes CMS uses and changes in the practitioner roster will affect each quarter's attribution differently, depending on the portion of that year that is in the lookback period for a quarter. For example, adding G0506 (assessment/care planning for patients requiring CCM services) as a CCM service starting on January 1, 2017, affected quarters from the second quarter of 2017 onward, since the second quarter of 2017 is the first quarter that contains 2017 in its lookback period.

5.B.4. How does attribution differ between the CPC+ evaluation and CMS payment?

Our attribution method for the evaluation identifies Medicare beneficiaries assigned to any practice each quarter using roughly the same claims-based attribution algorithm that CMS uses to attribute beneficiaries for CPC+ payments. However, our attribution approach for the evaluation differs from CMS's attribution approach in four key ways, described below.

A. The evaluation practitioner rosters come from IQVIA data for all practices (including CPC+ practices)

For payment attribution, CMS uses CPC+ practitioner rosters (lists of participating practitioners that practices participating in CPC+ submit to CMS) to determine the composition of CPC+ practices and their NPIs and TINs. However, analogous information about practice composition and TINs is not available for comparison practices. Therefore, to maintain consistency in identifying practice composition across CPC+ and comparison practices for the purposes of the evaluation, we use IQVIA's roster to obtain information on NPIs affiliated with a practice. Also, for both CPC+ and comparison practices, we assign TINs to each practice using an algorithm that picks the TIN that was most frequently billed in Medicare claims for primary care services by the NPIs at that practice.

Because we use IQVIA practitioner rosters for all practices, we group non-CPC+ practitioners into primary care practices, whereas payment attribution generally defines non-CPC+ practices as individual practitioners using single TIN-NPI or CCN-NPI combinations (because information regarding how they are grouped as actual practices is not available). The exception is that payment attribution defines practices that applied for CPC+ but were not accepted for CPC+ as practice sites using the practices' application rosters. The evaluation approach allows all non-CPC+ primary care practices in the frame, as well as any individual primary care practitioners not identified in IQVIA data, to compete with CPC+

[18] The number of attributed beneficiaries in the CPC+ and comparison groups changed minimally. For example, for 2019 Q2, the number of beneficiaries attributed to CPC+ practices decreased slightly from 1,817,130 for the fourth annual report, to 1,813,991 for the final report, or by 0.2 percent.

practices for beneficiaries. This process results in attributing fewer beneficiaries to CPC+ practices than the payment attribution process but likely leads to a more comparable attribution across CPC+ and non-CPC+ practices, because non-CPC+ practices compete for beneficiaries on equal footing with CPC+ practices.

B. The evaluation approach applies fewer restrictions to our definition of an attribution-eligible Medicare beneficiary

In CMS's payment attribution methodology, CMS excludes from attribution beneficiaries who: (1) have end-stage renal disease (ESRD) or are enrolled in hospice when they are first attributed (although beneficiaries with ESRD or hospice enrollment can be attributed if they were attributed to a CPC+ practice in an earlier quarter), (2) are in a long-term care institution, and (3) are enrolled in any other program that includes a Medicare FFS shared savings opportunity, except the Medicare Shared Savings Program (SSP).[19] However, for the evaluation, we do not apply any of these three exclusions in identifying attributed beneficiaries, because CMS expects CPC+ to affect all beneficiaries attributed to the practice, not just those for whom CMS calculates payments. In other words, for the evaluation, we want to assess impacts on all beneficiaries who received the plurality of their care from a CPC+ practice relative to similar beneficiaries attributed to comparison practices. Therefore, we think it is appropriate to apply only the eligibility criteria that pertain to the observability of the beneficiary's outcomes in Medicare FFS claims. CMS applies the same eligibility criteria in identifying attributed beneficiaries for payments, although the timing of these checks differs, as we describe below.

C. The evaluation's two-year lookback period begins immediately prior to the start of the quarter

For payment attribution, CMS uses a two-year claims lookback period that ends three months before the start of the quarter, because CMS needs the list of attributed beneficiaries before the start of the quarter to calculate the care management fees and other CPC+ payments, such as the Comprehensive Primary Care Payment for beneficiaries attributed to each CPC+ practice. For the impact analysis, however, the three-month gap between the end of the lookback period and the beginning of the quarter is unnecessary. Our objective is to identify the appropriate sample of attributed beneficiaries in both CPC+ and comparison practices, without the need for calculating payments in real time. Therefore, the two-year claims lookback period for attribution in the impact analysis ends the day before the start of the quarter.

The difference in the claims lookback period also leads to a difference between CMS's approach and the evaluation in the timing of the above-mentioned Medicare FFS eligibility checks. Specifically, CMS checks for eligibility one month before the start of the quarter, and we apply these eligibility criteria at the beginning of the quarter. For example, beneficiaries had to meet all eligibility criteria on December 1, 2017, to be eligible for CMS's payment attribution in the first quarter of 2018 (January 1, 2018–March 30, 2018) but they needed to meet the Medicare FFS eligibility criteria as of January 1, 2018, to be attributed to the evaluation sample.

D. CMS adjusted its payment attribution methodology over time

Starting with the first quarter of 2018, CMS included the Annual Wellness Visit and Welcome to Medicare visit criteria in its payment attribution process. Although we included this change in our

[19] During 2017 through 2021, examples of the excluded programs included Next Generation ACO, Comprehensive ESRD Care, the Financial Alignment Demonstration, and the Independence at Home Practice Demonstration.

attribution algorithm starting in the first quarter of 2018, it resulted in an additional discrepancy between the evaluation attribution for the fourth quarter of 2017 and payment attribution for the first quarter of 2018, the two quarters with identical claims lookback under each approach. Our attribution for 2017 Quarter 4 (Q4) covers the same lookback period as CMS's payment attribution for 2018 Q1. Because we do not include the Annual Wellness Visit criterion for the 2017 quarters, this could result in additional differences in attribution results between the evaluation sample for 2017 Q4 and the payment sample for 2018 Q1, the two quarters with identical claims lookback periods under each attribution algorithm.

Starting with the first quarter of 2019, CMS included an additional criterion based on voluntary assignment in its attribution process, as follows:

- If the beneficiary voluntarily attests that an eligible practitioner is the beneficiary's primary care physician, attribute the beneficiary to that practitioner's practice.
- For remaining beneficiaries, if the most recent primary care service was a CCM-service, attribute beneficiaries to the practice with the most recent CCM-related billing.
- Attribute remaining beneficiaries to the practice with the most recent Annual Wellness Visits or Welcome to Medicare Visits.
- Attribute all remaining beneficiaries to practices on the basis of the plurality of eligible primary care visits.

Because we do not include the voluntary assignment criterion, this could have resulted in additional differences between the evaluation and payment samples in quarters 2018 Q4 to 2021 Q4.[20] However, our preliminary analysis indicates that the extent of this additional discrepancy is very small, as fewer than half of one percent of beneficiaries voluntarily attest to a practitioner. We are unable to replicate the voluntary assignment criterion for the comparison group, so we do not include it in our attribution process for CPC+ or comparison practices.

Starting with the first quarter of 2021, CMS allowed beneficiaries attributed to SSP to also be attributed to CPC+ practices only if they are attributed to the SSP ACO that the CPC+ practice is affiliated with. We do not incorporate this change in our attribution process for two reasons. First, the evaluation attribution only applies the eligibility criteria relevant to the "observability" of the beneficiary's outcomes in Medicare FFS claims (for example, we require the beneficiary to be alive and enrolled in both Medicare Part A and Part B at the start of the quarter), and this change does not affect observability. Second, CMS adjusted the payment attribution to make it consistent with the method used in CMS's other primary care initiative, Primary Care First (PCF), which is not necessary for the CPC+ evaluation. This difference resulted in a slight decrease in the percentage of beneficiaries in our evaluation sample who are also in the payment sample from 2020 Q4 onwards, because more beneficiaries are considered ineligible during the lookback period in payment attribution.

In addition, in 2021, CMS started using a combined practitioner roster for CPC+ and PCF. Consequently, PCF practitioners, who were not previously on CPC+ rosters and thus treated as single primary care NPIs, now compete with CPC+ practices as a bigger "practice," which resulted in fewer beneficiaries attributed

[20] We compare 2018 Q4 of the evaluation attribution sample and 2019 Q1 of the payment attribution sample because they cover the same lookback period. Therefore, including voluntary assignment to payment attribution in 2019 Q1 impacts the overlap between the evaluation's sample for 2018 Q4 as well.

to CPC+ practices in payment attribution. Because the source data for a practitioner roster used for evaluation attribution (i.e., IQVIA data) did not experience the same change between 2021 and previous years, this change further reduced the percentage of beneficiaries in the evaluation sample who are also in the payment sample in 2021 Q3 and Q4, the two quarters in which the 2021 roster was used in the 24-month lookback period in payment attribution.

The similarities and differences between CMS's approach and the evaluation's approach for beneficiary attribution are summarized in Table 5.B.4.

Table 5.B.4. Similarities and differences between beneficiary attribution for payment versus evaluation through 2021

	Payment attribution	Evaluation attribution
Similarities between payment and evaluation attribution processes		
Frequency of attribution	Quarterly	Same as payment attribution.
Observability criteria for beneficiary eligibility	Be enrolled in Medicare Part A and Part B. Not be covered under a Medicare Advantage or other Medicare health plan. Not be incarcerated. Be alive.	Same as payment attribution.
Criteria used to identify eligible services for attribution	Evaluation and management HCPCS codes.	Same as payment attribution.
Attribution algorithm for 2017 quarters	If the most recent primary care service was a CCM service, attribute beneficiaries to the practice with most recent CCM-related billing. Attribute all remaining beneficiaries to practices on the basis of the plurality of eligible primary care visits.	Same as payment attribution.
Attribution algorithm for 2018 quarters	If the most recent primary care service was a CCM service, attribute beneficiaries to the practice with most recent CCM-related billing. If the most recent visit was not a CCM service, and the beneficiary had an Annual Wellness Visit or a Welcome to Medicare visit, attribute the beneficiary to the practice that had most recent Annual Wellness Visit or Welcome to Medicare visit. Attribute all remaining beneficiaries to practices on the basis of the plurality of eligible primary care visits.	Same as payment attribution.
Differences between payment and evaluation attribution processes		
Attribution algorithm for 2019 and 2021 quarters	If beneficiaries voluntarily attest that an eligible practitioner is their primary care physician, attribute the beneficiaries to that practitioner's practice. For the remaining beneficiaries, if the most recent primary care service was a CCM service, attribute the beneficiaries to the practice with the most recent CCM-related billing. If the most recent visit was not a CCM service, and the beneficiaries had an Annual Wellness Visit or a Welcome to Medicare visit, attribute the beneficiaries to the practice that had the most recent Annual Wellness Visit or Welcome to Medicare visit. Attribute all remaining beneficiaries to practices on the basis of the plurality of eligible primary care visits.	Same as payment attribution, except we cannot approximate voluntary attestation.

Table 5.B.4 *(continued)*

	Payment attribution	Evaluation attribution
Time period for conducting attribution	Intervention quarters.	Baseline and intervention quarters.
Source for roster of practices and their practitioners	CPC+ practitioner rosters (CPC+/PCF practitioner roster in 2021).	IQVIA's SK&A rosters from baseline through PY 2. IQVIA's OneKey rosters from PY 3 through PY 5.
Source for TINs	CPC+ practitioner rosters (CPC+/PCF practitioner roster in 2021).	TIN assignment process based on claims.
Practices/practitioners with whom CPC+ practices compete for beneficiaries	Practices rejected from CPC+ and single primary care NPIs not on CPC+ rosters (and not on PCF roster in 2021).	All primary care practices from IQVIA roster and single primary care NPIs not on IQVIA roster.
Additional criteria for beneficiary eligibility	Cannot have end-stage renal disease and cannot be enrolled in hospice when they are first attributed. Cannot be in a long-term care institution. Cannot be enrolled in program that includes a Medicare FFS shared savings opportunity, except SSP.	Can have end-stage renal disease or be enrolled in hospice. Can be in a long-term care institution. Can be enrolled in program that includes a Medicare FFS shared savings opportunity.
Time frame for evaluating eligibility criteria	Three months before the start of the quarter for 2017 Q1–2017 Q2. Otherwise, one month before start of quarter.	Day of the start of quarter.
Lookback period for claims used in quarter's attribution process	Two-year period that ends three months before the start of the quarter.	Two-year period that ends immediately before the start of the quarter.
Tie-breaker to determine the practice with the most visits among those that have the same number of visits and same date of most recent visit	Preference given to CPC+ practices over all other practices and NPIs.	No preference given to CPC+ practices relative to comparison practices (all practices on IQVIA roster are given preference over all other single primary care NPIs not on IQVIA roster).

CCM = Chronic Care Management; FFS = fee-for-service; HCPCS = Healthcare Common Procedure Coding System; NPI = National Provider Identifier; PCF = Primary Care First; Q = quarter; SSP = Medicare Shared Savings Program; TIN = Tax Identification Number.

5.B.5. How similar are the evaluation attribution samples to CMS's payment attribution samples?

Given the differences in attribution methodology between CPC+ payment and the CPC+ evaluation, the evaluation is unlikely to attribute 100 percent of the same beneficiaries to CPC+ practices as CMS does for payment attribution. The biggest concern is the difference between using the practitioner rosters and using IQVIA data and TIN assignment—because including different sets of practitioners within practices could lead to large differences in the beneficiaries attributed to the practices.

If there are large differences between the payment attribution sample and the evaluation sample, that could mean that the beneficiaries in our evaluation sample are not actually under the care of CPC+

practices—and thus they are not expected to be impacted by CPC+.[21] This would lead to attenuation in the impact estimates.

Therefore, it is important to track how well the Medicare beneficiary sample used in the evaluation and the Medicare beneficiary sample used by CMS for payments to CPC+ practices align.

To do this, we implement the following analyses.

First, we calculate the overlap of practitioners assigned to CPC+ practices based on the practitioner roster submitted to CMS and those on the practitioner rosters we develop using data purchased each year from IQVIA to support patient attribution for the evaluation. We used data from IQVIA's SK&A database for the baseline period and the first two years of CPC+, and data from IQVIA's OneKey database starting in PY 3. When we construct our master practice-practitioner file, we use the practice location and practice address to identify practices participating in CPC+ in the data received from IQVIA. However, even though the two data sources might indicate the same practice by practice name and location, there might be important differences in the list of practitioners between the two rosters that would affect beneficiary attribution.

To check the overlap of practitioners across the two rosters, we merge CPC+ program data with IQVIA data by practitioner NPI and report (1) the percentage of practitioners in CPC+ rosters who were found in the IQVIA rosters of these practices and (2) the percentage of practitioners in IQVIA rosters for these practices who were found in the CPC+ rosters. We limit CPC+ rosters to practitioners marked as actively participating in CPC+ to remove practitioners who may have moved to another location. In Table 5.B.5, we compare CPC+ practitioner rosters to IQVIA practitioner rosters at six time points: one month before CPC+ began (December 2016), month 12 of CPC+ (December 2017), month 24 of CPC+ (December 2018), month 36 of CPC+ (December 2019), month 48 of CPC+ (December 2020), and month 60 of CPC+ (December 2021). We found 74.3 to 81.0 percent of active practitioners in the CPC+ rosters appeared in the SK&A rosters (Table 5.B.5) between baseline and PY 2 of CPC+, with the percentage overlap declining over time. IQVIA's switch to using the OneKey database for the rosters improved the overlap rate to 85.0 to 85.3 percent for PY 3 through PY 5 of CPC+.[22]

The percentage of IQVIA practitioners found as active practitioners in CPC+ rosters declined over time from 82.5 percent at baseline to 58.8 percent by PY 5. This decline over time is partly due to practices withdrawing or being terminated from CPC+. Those practices and their practitioners are removed (marked inactive) from the CPC+ roster but remain part of the intervention sample given the evaluation's intent-to-treat approach.

Note that we do not see a strong decline in the percentage of beneficiaries in the evaluation sample who are also in the payment sample (Table 5.B.6). It remains above 89 percent throughout the intervention period. This makes us less concerned about the decline in the percentage of practitioners in the IQVIA

[21] It is also possible that the CPC+ payment sample might include beneficiaries for whom the practices are not truly responsible; however, once beneficiaries become attributed to a CPC+ practice, that practice has an incentive to make sure they receive high quality care.

[22] We expect that this increase in number of practitioners in the CMS roster who are found in the IQVIA rosters is because the OneKey data capture more practitioners by bringing in data from administrative sources, whereas SK&A relied primarily on phone verification to collect practitioner data.

practitioner roster who are also in the CPC+ roster, because the beneficiary overlap is what matters for our beneficiary-level impact analysis.

Table 5.B.5. CMS and IQVIA primary care practitioner roster comparison

Compared rosters	Before CPC+ began (Baseline)	One year after CPC+ began (PY 1)	Two years after CPC+ began (PY 2)	Three years after CPC+ began (PY 3)	Four years after CPC+ began (PY 4)	Five years after CPC+ began (PY 5)
Number of practices	2,865[a]	2,888	2,888	2,888	2,888	2,888
Unique primary care practitioners						
Number of active practitioners in CPC+ roster	12,950	13,342	13,182	13,049	12,962	12,200
Number of practitioners in IQVIA roster	12,712	13,299	13,820	16,844	17,028	17,700
Percentage of active practitioners in the CPC+ roster also in the IQVIA roster	81.0	78.1	74.3	85.3	85.0	85.3
Percentage of practitioners in the IQVIA roster also active in the CPC+ roster	82.5	78.4	70.9	66.6	64.7	58.8

Notes: All duplicate NPIs were removed from both rosters. The baseline comparison is based on December 2016 data; the PY 1 comparison uses December 2017 data; the PY 2 comparison uses December 2018 data; the PY 3 comparison uses December 2019 data; the PY 4 comparison uses December 2020 data, and the PY 5 comparison uses December 2021 data. Baseline, PY 1, and PY 2 IQVIA rosters are based on SK&A data, while PY 3, PY 4, and PY 5 IQVIA rosters are based on OneKey data. The IQVIA practitioner roster is restricted to primary care practitioners; we identified a practitioner as primary care using primary and secondary taxonomy codes in the NPPES and specialty information included on Medicare claims over a 12-month lookback period. We do not restrict the CMS rosters since they should already be restricted to primary care practitioners. The IQVIA data rows include 148 practices that we were unable to find in the IQVIA data, but for which we supplemented the IQVIA data with CPC+ roster data.

[a] We were unable to find either SK&A or CMS's CPC+ roster information for 23 practices at baseline. Once the intervention began, we added these practices using the CMS roster from February 2017.

NPI = National Provider Identifier; NPPES = National Plan & Provider Enumeration System; PY = Program Year.

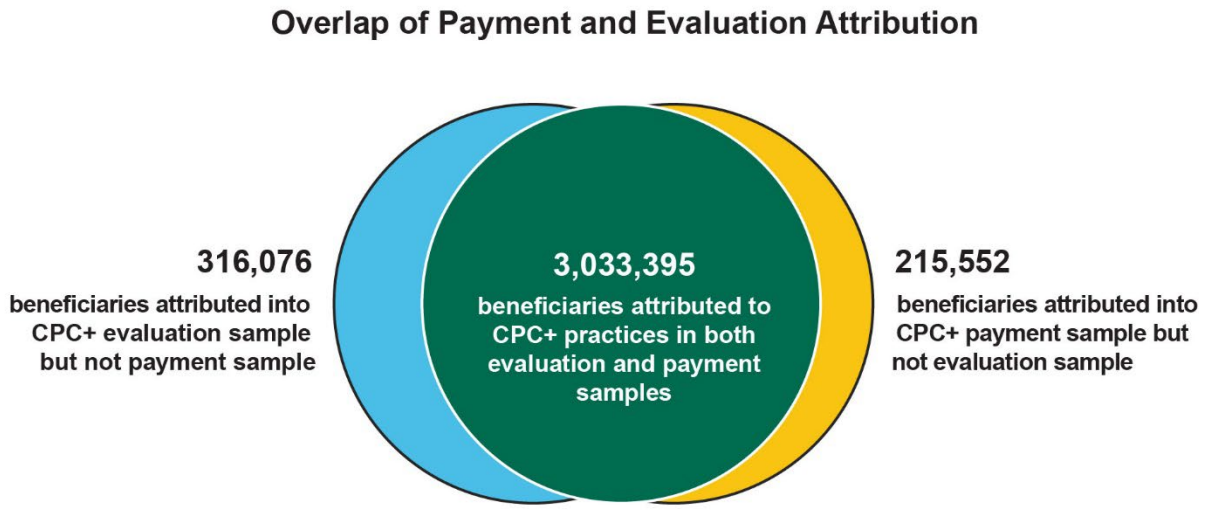
Second, we calculate the overlap in beneficiaries attributed to CPC+ practices in the payment and evaluation samples. Due to the differences in the lookback period for a specific calendar quarter (see Subsection C above in Section 5.B.4), we compare each evaluation sample to the subsequent quarter's payment sample. For example, we compare the evaluation sample from 2017 Q1 (January–March 2017) to the payment sample from 2017 Q2 (April–June 2017). This ensures we are comparing attribution from quarters that use the same lookback period in the payment and evaluation samples. In addition to all the intervention quarters, CMS only ran payment attribution for baseline quarters 2016 Q1 and Q4, so we

cannot compare our attribution for 2016 Q2 and Q3 to the equivalent payment attribution sample. Further, we cannot compare the 2021 Q4 evaluation sample to the subsequent quarter's payment sample because payment attribution ended in 2021 Q4.

We found substantial overlap between the samples of beneficiaries ever attributed to CPC+ practices by CMS and by the evaluation over the five years of the intervention. As we show in Figure 5.B.1, 3,033,395 Medicare beneficiaries were ever attributed to CPC+ practices in both the evaluation sample and the sample CMS used for payment; 215,552 beneficiaries were ever attributed to the CPC+ payment sample but never to the evaluation sample; and 316,076 were ever attributed to the CPC+ evaluation sample but never to the payment sample. More specifically, Table 5.B.6 shows that 87 percent or more of the beneficiaries attributed to 2017 Starter CPC+ practices in our evaluation sample for the first 19 CPC+ quarters were also attributed to the payment attribution sample in the equivalent quarter. Also, 86 to 90 percent of beneficiaries attributed to the payment attribution sample by CMS each quarter were also attributed to CPC+ practices for the evaluation in the equivalent quarter.

Third, using CMS's payment eligibility criteria, we calculate the number of beneficiaries we attribute to CPC+ practices who would have been eligible for payment attribution. This involves additionally limiting the sample to beneficiaries who are not receiving hospice, do not have ESRD, are not institutionalized, and are not enrolled in any other program that includes a Medicare FFS shared savings opportunity, except SSP. Table 5.B.6, column 5, reports the number of beneficiaries in the evaluation sample for each quarter, and column 6 reports the number of beneficiaries in the evaluation sample under CMS's payment eligibility rules. This difference is approximately 40,000 or 2.5 percent of the evaluation sample in a given quarter.

Figure 5.B.1. Attribution of Medicare FFS beneficiaries during PY 1 through PY 5



Source: Comparison of attributed Medicare FFS beneficiaries in Mathematica's evaluation sample for the 1st through the 19th program quarters (January 2017–September 2021) and those in CMS's payment sample for the 2nd through the 20th program quarters (April 2017–December 2021), which used the same set of two-year lookback periods. We used Medicare FFS beneficiary lists provided by CMS to define the payment sample.

FFS = fee-for-service; PY = Program Year.

Table 5.B.6. Beneficiaries attributed to 2017 Starter CPC+ practices, by quarter

Mathematica attribution quarter	Comparison to payment quarter	Beneficiaries in both payment and evaluation samples	Beneficiaries in payment sample	Beneficiaries in evaluation sample	Beneficiaries in evaluation sample under payment eligibility rules	Percentage of beneficiaries in payment sample who are in evaluation sample	Percentage of beneficiaries in evaluation sample who are in payment sample
2016 Q1	2016 Q2	1,489,022	1,655,920	1,651,432	1,609,642	90%	90%
2016 Q2	NA	NA	NA	1,720,593	1,680,865	NA	NA
2016 Q3	NA	NA	NA	1,773,509	1,734,138	NA	NA
2016 Q4	2017 Q1	1,638,668	1,820,621	1,810,383	1,770,994	90%	91%
2017 Q1	2017 Q2	1,607,043	1,795,086	1,767,439	1,723,511	90%	91%
2017 Q2	2017 Q3	1,647,250	1,847,515	1,795,295	1,755,187	89%	92%
2017 Q3	2017 Q4	1,676,565	1,894,700	1,816,139	1,776,977	88%	92%
2017 Q4	2018 Q1[a]	1,668,424	1,937,859	1,833,634	1,794,859	86%	91%
2018 Q1	2018 Q2	1,692,514	1,907,212	1,826,664	1,784,426	89%	93%
2018 Q2	2018 Q3	1,707,502	1,930,223	1,844,365	1,803,384	88%	93%
2018 Q3	2018 Q4	1,716,965	1,950,103	1,856,681	1,815,803	88%	92%
2018 Q4	2019 Q1[b]	1,711,262	1,955,435	1,865,477	1,824,614	88%	92%
2019 Q1	2019 Q2	1,644,503	1,897,910	1,783,561	1,744,690	87%	92%
2019 Q2	2019 Q3	1,664,762	1,915,740	1,813,991	1,775,816	87%	92%
2019 Q3	2019 Q4	1,681,130	1,922,162	1,840,587	1,802,321	87%	91%
2019 Q4	2020 Q1	1,686,634	1,917,936	1,864,457	1,826,189	88%	90%
2020 Q1	2020 Q2	1,635,024	1,850,709	1,803,349	1,766,142	88%	91%
2020 Q2	2020 Q3	1,645,443	1,855,136	1,816,773	1,780,584	89%	91%
2020 Q3	2020 Q4	1,638,455	1,843,779	1,812,987	1,777,652	89%	90%
2020 Q4	2021 Q1[c]	1,625,026	1,822,561	1,827,529	1,793,246	89%	89%
2021 Q1	2021 Q2	1,561,097	1,756,706	1,748,073	1,699,962	89%	89%
2021 Q2	2021 Q3[d]	1,524,721	1,713,295	1,754,666	1,668,958	89%	87%
2021 Q3	2021 Q4	1,527,321	1,714,201	1,762,951	1,677,465	89%	87%
2021 Q4	NA	NA	NA	1,765,810	1,680,493	NA	NA

Source: Comparison of attributed Medicare FFS beneficiaries in Mathematica's evaluation sample for the 1st through the 19th program quarters (January 2017–September 2021) and those in CMS's payment sample for the 2nd through the 20th program quarters (April 2017–December 2021), which used the same set of two-year lookback periods. We were not able to compare the 2021 Q4 evaluation sample to the subsequent quarter's payment sample because payment attribution ended in 2021 Q4. We used Medicare FFS beneficiary lists provided by CMS to define the payment sample.

[a] In 2018, CMS changed its attribution rules to prioritize practices in which beneficiaries had their most recent Annual Wellness Visit, which results in additional differences between the evaluation attribution for 2017 Q4 and the payment attribution for 2018 Q1, the two quarters with the same claims lookback period under each attribution algorithm. Starting in 2018 Q1, we incorporated this criterion into the evaluation attribution rules as well.

[b] In 2019, CMS changed its attribution rules to prioritize practices in which beneficiaries had voluntarily assigned themselves, which results in additional differences in attribution.

[c] In 2021, CMS changed its attribution rules to allow beneficiaries attributed to SSP to also be attributed to CPC+ practices only if they are attributed to the SSP ACO that the CPC+ practice is affiliated with, which results in additional differences in attribution.

[d] In 2021, CMS started using a combined practitioner roster for CPC+ and PCF, which results in fewer beneficiaries attributed to CPC+ practices in payment attribution for 2021 Q3 and Q4 (and also a lower percentage of beneficiaries in the evaluation sample who are in the payment sample), the two quarters in which the 2021 roster was used in the 24-month lookback period. This is because PCF practitioners, who were not previously on CPC+ rosters and thus treated as single primary care NPIs, now compete with CPC+ practices as a bigger "practice" in payment attribution. The source data for a practitioner roster used for evaluation attribution (i.e., IQVIA data) did not experience the same change between 2021 and previous years.

ACO = Accountable Care Organization; FFS = fee-for-service; NA = not available; NPI = National Provider Identifier; PCF = Primary Care First; Q = quarter; SSP = Medicare Shared Savings Program.

5.C. Specification of measures used in the Medicare impact analysis

In this Appendix, we define the key measures used in this report that are based on Medicare claims and enrollment information. First, we define and discuss the Medicare claims-based outcome measures used in the impact analysis. Next, we describe non-outcome measures based on Medicare claims and enrollment data that we used as control variables in the regression analysis or for other analyses. We also describe updates or changes to outcomes since the fourth annual report. All updates or changes are applied to all measurement years.

5.C.1. Medicare claims-based outcome measures

Table 5.C.1 summarizes the outcome measures we used in the annual impact analysis in this report. We classified the claims-based outcome measures into groups by Medicare expenditures, service utilization, and four of the five CPC+ functions (improvements in planned care and population health, continuity of care, comprehensiveness of care, and patient and caregiver engagement). Relative to the fourth annual report, we added new outcome measures, which are listed along with their motivation in Table 5.C.2.

For each outcome, we show the hypothesized direction of impact in Table 5.C.1. For some measures, the expected direction of effect is indeterminate, because there are multiple mechanisms that could either increase or decrease the outcome, and it is not clear which mechanism would or should outweigh the other. For example, ambulatory specialist visits could increase or decrease, depending on the extent to which more effective care management and follow-up after hospitalizations by CPC+ practices reduce the need for specialist visits or result in more referrals to specialists.

Table 5.C.1. Medicare claims-based outcome measures for the final report for the independent evaluation of CPC+

	Hypothesized direction of impact
Medicare Parts A and B expenditures (PBPM)	
Excluding enhanced payments[a]	↓
Including CPC+ CMFs[b]	↓ or →
Including CPC+ CMFs, PBIPs, and shared savings payments to SSP ACOs[b]	↓ or →
Monthly Medicare expenditures by service category (PBPM)[c]	
Inpatient: Expenditures for both acute inpatient care (short-stay acute hospitals and CAHs) and non-acute inpatient care (e.g., inpatient rehabilitation services, psychiatric hospital services, etc.)	↓
Expenditures for acute inpatient care[d]	↓
Expenditures for acute surgical hospitalizations	↓
Expenditures for acute medical (i.e., non-surgical) hospitalizations	↓
Expenditures for inpatient rehabilitation facilities[e]	↓
Post-acute care expenditures (post-acute home health, long-term care, skilled nursing facility, and inpatient rehabilitation)	↓
Expenditures for acute inpatient care and post-acute care	↓
Outpatient: Outpatient facility expenditures including those for ED visits (including observation stays), and other outpatient services (e.g., outpatient surgery, imaging, outpatient rehabilitation, and services provided by RHCs and FQHCs)	↓
Expenditures for outpatient ED visits, including observation stays[f]	↓
Physician and nonphysician Part B noninstitutional services: Expenditures including physician services and other services provided by ambulance providers, independent clinical laboratories, and freestanding ambulatory surgical centers[g]	↑ or ↓
Ambulatory visits with primary care practitioners: Expenditures for visits with a primary care practitioner in noninstitutional settings (e.g., office, home, hospital outpatient department, FQHC, RHC, CAH, etc.)	↑ or ↓
Proportion of expenditures for ambulatory visits with primary care practitioners that are non-face-to-face[h]	↑ or →
Expenditures for ambulatory visits with primary care practitioners at assigned practice[i]	↑ or ↓
Expenditures for ambulatory visits with primary care practitioners at non-assigned practice[i]	↑ or ↓
Ambulatory visits with specialists: Expenditures for visits with a specialist in noninstitutional settings: (e.g., office, home, hospital outpatient department, FQHC, RHC, CAH, etc.)	↑ or ↓
Laboratory expenditures[j]	↑ or ↓
Imaging expenditures[j]	↑ or ↓
Skilled nursing facility expenditures	↓
Home health expenditures	↑ or ↓
Hospice: Expenditures for hospice providers in both institutional and home settings	↑
Durable medical equipment: Expenditures for DME, such as wheelchairs, home oxygen, and home hospital beds	↑ or ↓

Table 5.C.1. ***(continued)***

	Hypothesized direction of impact
Annualized service use (per 1,000 beneficiaries per year)	
Number of hospitalizations (short-stay acute care and CAHs)	↓
Number of acute surgical hospitalizations	↓
Number of acute surgical hospitalizations with a major complication or comorbidity	↓
Number of acute surgical hospitalizations with a complication or comorbidity	↓
Number of acute surgical hospitalizations without any complication or comorbidity	↓
Number of acute medical (i.e., non-surgical) hospitalizations	↓
Number of acute medical hospitalizations with a major complication or comorbidity	↓
Number of acute medical hospitalizations with a complication or comorbidity	↓
Number of acute medical hospitalizations without any complication or comorbidity	↓
Total number of ED visits, including observation stays (outpatient ED visits and ED visits resulting in a hospitalization)[k]	↓
Number of outpatient ED visits including observation stays (one overall measure and one measure that excludes services primarily for a COVID-19 diagnosis or a respiratory condition potentially caused by COVID-19)	↓
Number of primary care substitutable outpatient ED visits[l]	↓
Number of potentially primary care preventable outpatient ED visits[l]	↓
Total number of UCC visits (one overall measure and one measure that excludes services primarily for a COVID-19 diagnosis or a respiratory condition potentially caused by COVID-19)	↓
Number of primary care substitutable UCC visits	↓
Number of ambulatory primary care visits (including to FQHCs, RHCs, and CAHs)[m]	↑ or ↓
Proportion of ambulatory primary care visits that are non-face-to-face[h]	↑ or →
Number of ambulatory specialist visits (including to FQHCs, RHCs, and CAHs)[m]	↑ or ↓
Planned care and population health (annualized)	
Among Medicare FFS beneficiaries ages 18–75 with diabetes, percentage who received[n]	
Hemoglobin A1c test	↑
Retinal eye exam	↑
Medical attention for nephropathy	↑
All three tests (HbA1c test, retinal eye exam, and medical attention for nephropathy)	↑
None of the three tests	↓
Among female Medicare FFS beneficiaries ages 52–74, percentage who received:[o]	
Breast cancer screening	↑
Among Medicare FFS beneficiaries ages 21 and older:[p]	
Percentage of beneficiaries with cardiovascular disease who were prescribed and filled statin therapy	↑
Among Medicare FFS beneficiaries ages 18 and older:[p]	
Percentage of beneficiaries on diabetes medications with proportion of days covered by medication > 80%	↑
Percentage of beneficiaries on renin-angiotensin system antagonists with proportion of days covered by medication > 80%	↑
Percentage of beneficiaries on statins with proportion of days covered by medication > 80%	↑
Percentage of beneficiaries with both coronary artery disease (CAD) and diabetes who were prescribed and filled angiotensin-converting enzyme (ACE) inhibitors or angiotensin receptor blockers (ARB) therapy	↑

Table 5.C.1. ***(continued)***

	Hypothesized direction of impact
Continuity of care	
Percentage of primary care ambulatory visits provided at a beneficiary's assigned practice[i,q]	↑ or ↓
Among beneficiaries with qualifying ambulatory visits in the measurement year:[r]	
Across all PCPs and specialists providing care to a patient, where each practitioner in the beneficiary's assigned practice is treated separately:	
Percentage of visits with the usual provider of care (UPC)[s]	↑
Reversed Bice-Boxerman Index (rBBI)[t]	↓
Across all PCPs and specialists providing care to a patient, where all practitioners in the beneficiary's assigned practice are treated as a single practitioner:	
Percentage of visits with the UPC[s]	↑
Reversed Bice-Boxerman Index (rBBI)[t]	↓
Comprehensiveness of care (measured at the NPI level)[u]	
Involvement in patient conditions[v]	↑
New problem management[w]	↑
Range of services provided by primary care physicians[x]	↑
Patient and caregiver engagement	
Hospice service use:	
Percentage of beneficiaries receiving hospice services	↑
Days of hospice use for beneficiaries receiving hospice services in the measurement year[y]	↑
Days of hospice use for all beneficiaries in the measurement year	↑
Other Quality of Care	
Percentage of index acute care hospital discharges that were followed by an unplanned readmission within 30 days[z]	↓
Percentage of index acute care hospital discharges that were followed by an unplanned acute care hospitalization or ED visit (including observation stays) within 30 days[aa]	↓
Percentage of index ED (including observation stay) discharges that were followed by an unplanned acute care hospitalization or ED visit (including observation stays) within 30 days[aa]	↓
Among Medicare FFS beneficiaries ages 65 and older, percentage who received:[bb]	
Two or more prescriptions for high risk medications in the same medication class	↓
Long-term opioid use[cc]	↓
Potential opioid overuse[cc, dd]	↓
Annualized number of low-value services per 1,000 beneficiaries[ee]	↓

Note: For the Medicare expenditures and service utilization measures, services and costs are only counted during months that a beneficiary is enrolled in FFS Parts A and B with Medicare as the primary payer, not enrolled in a health maintenance organization (HMO), and alive. For other measures, such as those for planned care and population health and other quality of care outcomes, we follow the guidelines of the measure stewards and note any deviations from those specifications. In general, for the quality-of-care outcomes and the continuity of care measures, specific criteria are used to identify beneficiaries eligible for each measure and continuous enrollment in Medicare Parts A and B (also in Part D, for measures based on prescription drug use) is required during the measurement period. We provide details on these restrictions in the description of each measure in the sections below.

[a] Expenditures for Part A and Part B services in PY 3 and PY 4 include QPP payment adjustments, based on practitioner performance two years before. They are applicable for both CPC+ and comparison practices. The adjustments are composed of (1) MIPS adjustments, which are applied directly to physician and outpatient claims (as a percentage of the charges on the claims), and (2) lump sum incentive payments to eligible practitioners who participated in Advanced APMs in 2017 and 2018 (calculated based on 2018 and 2019 claims for these practitioners, respectively). For Track 2 practices, Medicare Part A and B expenditures *without* enhanced payments include the base CPCPs, but not the 10 percent comprehensiveness supplement. We include CPCPs in Part B spending because Track 2 practices agreed to lower Part B payment for evaluation and management (E&M) services in exchange for CPCPs.

[b] For Track 2 practices, Medicare Parts A and B expenditures *with* enhanced payments include the base CPCPs, as well as the 10 percent comprehensiveness supplement.

Table 5.C.1. ***(continued)***

[c] The sum of expenditures by service category does not equal the total expenditures for Part A and Part B services without enhanced payments because the total expenditures include lump sum incentive payments that are not applied at the claim level and are instead paid out directly to eligible practitioners who participated in Advanced APMs in 2017 and 2018.

[d] Acute inpatient care includes short-stay acute hospital admissions and admissions to CAHs.

[e] Expenditures for non-acute hospital admissions other than those for inpatient rehabilitation, such as psychiatric hospital admissions, are included in inpatient expenditures but not shown separately in the report.

[f] Expenditures, with QPP payment adjustments, for outpatient ED visits include professional and facility fees, as well as payments for observation stays. Although these expenditures are shown under outpatient expenditures, they include professional fees, which are part of expenditures for physician and nonphysician Part B noninstitutional services.

[g] Expenditures, with QPP payment adjustments, for Part B noninstitutional services include expenditures for (1) ambulatory primary care visits, (2) ambulatory specialist visits, and (3) non-ambulatory physician visits, as well as services provided by other noninstitutional providers (the third category is not shown separately).

[h] Ambulatory visits are identified as face-to-face or non-face-to-face based on procedure codes, telehealth modifiers, and place of service (carrier file only) on Medicare claims. Visits such as telephone and online assessment and management and E&M are included in the non-face-to-face measure, making it broader than CMS's definition of "telehealth" visits. The sum of the face-to-face and non-face-to-face visits with primary care practitioners equals the total ambulatory visits. Since the number of non-face-to-face visits were infrequent before the COVID-19 pandemic, we examined non-face-to-face visits and expenditures on such visits in PY 4 and PY 5 only.

[i] We define the assigned practice for the baseline period as the first practice to which a beneficiary was attributed during the baseline period, and the assigned practice for the intervention period as the first practice that the beneficiary was attributed to during the intervention period. Effects on this set of measures are ambiguous because CPC+ could increase the total number of visits as primary care practices offer more comprehensive services and, potentially, extend their office hours. Conversely, CPC+ could decrease in-person office visits by using other non-visit approaches for contacting patients (such as e-visits or secure messaging) or using non-billing care team members to deliver care. We particularly expect shifts to non-visit-based approaches among Track 2 practices, which are required to offer their patients at least one alternative to traditional office visits, in return for additional non-visit-based revenue in the form of the CPCP (and have their FFS amounts for those E&M services reduced).

[j] Laboratory and imaging services were identified in the carrier and outpatient files.

[k] Total ED visits include ED/observation stays that led to a hospitalization (including psychiatric hospitalizations).

[l] The sum of primary care substitutable outpatient ED visits and potentially primary care preventable outpatient ED visits is less than total outpatient ED visits because total outpatient ED visits include visits for other care needs, such as injuries, mental health, drug use, and alcohol use.

[m] Ambulatory visits with primary care practitioners and specialists include office-based visits and visits at home, as well as visits in other settings, such as FQHCs, RHCs, and CAHs.

[n] This measure requires that beneficiaries be continuously enrolled in Medicare FFS Parts A and B and not have hospice services during the measurement year.

[o] This measure requires that beneficiaries be continuously enrolled in Medicare FFS Parts A and B during the measurement year as well as the 27 months prior to October 1 of the measurement year and not have hospice services during the measurement year.

[p] This measure requires that beneficiaries be continuously enrolled in Medicare FFS Parts A and B as well as in Medicare Part D, and not use hospice services during the measurement year.

[q] Due to the intent-to-treat (ITT) approach for beneficiary assignment, we expect to see a decrease in visits to practitioners affiliated with the beneficiary's assigned practice over time for both CPC+ and comparison practices. This decline occurs because we continue to assign the beneficiary to the first practice the beneficiary was ever attributed to in the intervention period, regardless of whether the beneficiary continued to receive care at that practice.

[r] The continuity of care measures are calculated for beneficiaries who were in the ITT sample at the beginning of each measurement year, enrolled in Medicare FFS Parts A and B for the full year, and had qualifying ambulatory visits in that year. Qualifying ambulatory visits are office or other outpatient visits for (1) evaluation and management; (2) ophthalmological services: medical examination and evaluation; and (3) new enrollee and annual wellness visits.

[s] Beneficiaries must have one or more qualifying ambulatory visits to be included in the percentage of visits with the UPC measure.

[t] Beneficiaries must have four or more qualifying ambulatory visits to be included in the rBBI measure.

[u] NPIs are used to define the comprehensiveness of care at the practitioner level.

[v] For each NPI, this measure calculates the percentage of beneficiaries for whom the NPI was considered "most comprehensive" out of all beneficiaries the NPI saw in the year. "Most comprehensive" for this measure means that the NPI saw the patient for the largest share of their unique diagnosis codes. If two NPIs saw the patient for the largest share of their unique diagnosis codes, both NPIs are considered "most comprehensive" for the patient.

[w] Creates a score that indicates how often a primary care physician continues to treat a beneficiary's new condition versus referring the beneficiary (or the beneficiary self-referring) to a specialist or different provider.

[x] Creates a score (0–5) that counts the number of service categories for which that primary care practitioner (PCP) billed. The five service categories included in the measure are: immunization, behavioral or mental health counseling, treatment of minor lacerations, cryotherapy/skin excision, and joint injection.

[y] Calculated only for beneficiaries who had at least one day of hospice use during the measurement year.

[z] The readmissions outcome is per index discharge.

Table 5.C.1. ***(continued)***

[aa] There are two different unplanned acute care outcomes, depending on whether the index event was a hospital discharge or an ED discharge. Also, the definition of unplanned acute care is broad and consists of hospitalizations and ED visits, including observation stays.

[bb] This measure requires that beneficiaries be continuously enrolled in Medicare Parts A and B as well as in Medicare Part D, and not use hospice services during the measurement year.

[cc] To be included in the analysis of both long-term opioid use and potential overuse, a beneficiary had to: (1) be continuously enrolled in Medicare Parts A, B, and D throughout each measurement year or until death; and (2) have at least one opioid prescription during the measurement year. We further excluded beneficiaries for whom opioid use is appropriate: beneficiaries with a diagnosis of cancer during the measurement year or one year before, or a diagnosis of sickle cell disease or hospice use during the measurement year.

[dd] This measure is defined only among long-term users of opioids.

[ee] Mathematica recently updated the low-value services measure set, originally developed by Schwartz et al. (2014), to reflect ICD-10 diagnosis codes and the Restructured BETOS Classification System (RBCS) and by replacing the three low-value services that showed declining informativeness for current beneficiaries with three new services (Fleming et al. 2022).

ACO = Accountable Care Organization; APM = Alternative Payment Model; CAH = Critical Access Hospital; CMF = care management fee; CPCP = Comprehensive Primary Care Payment; DME = durable medical equipment; ED = emergency department; E&M = evaluation and management; FFS= fee-for-service; FQHC = Federally Qualified Health Center; MIPS = Merit-based Incentive Payment System; NPI = National Provider Identifier; PBIP = Performance-based Incentive Payment; PBPM = per beneficiary per month; PCP = primary care practitioner; QPP = Quality Payment Program; SSP = Medicare Shared Savings Program; RHC = Rural Health Clinic; UCC = urgent care center.

Table 5.C.2. Motivation for new CPC+ outcome measures

Outcomes by domain	Why is the outcome important to CPC+?
Medicare expenditure outcomes	
Monthly Medicare expenditures by service category (PBPM)	
Expenditures for acute surgical and medical hospitalizations, separately	• To further investigate which types of hospitalizations drove observed effects of CPC+ on expenditures for all-cause acute hospitalizations.
Post-acute care expenditures (post-acute home health care, long-term care, skilled nursing facility, rehabilitation)	• To further investigate the source of the observed effects of CPC+ on associated inpatient expenditures and utilization.
Expenditures for acute inpatient care and post-acute care	• To further investigate the source of the observed effects of CPC+ on inpatient and associated expenditures and utilization.
Laboratory expenditures	• To further investigate the source of the observed differential effects of CPC+ on expenditures of SSP and non-SSP practices.
Imaging expenditures	• To further investigate the source of the observed differential effects of CPC+ on expenditures of SSP and non-SSP practices.
Service use outcomes	
Number of acute surgical and medical hospitalizations, separately	• To further investigate which types of hospitalizations drove observed effects of CPC+ on all-cause acute hospitalizations.
Number of acute surgical hospitalizations, separately, by severity based on complication or comorbidity[a]	• To further investigate the source of any observed differential effects of CPC+ on acute surgical hospitalizations.
Number of acute medical hospitalizations, separately, by severity based on complication or comorbidity[a]	• To further investigate the source of any observed differential effects of CPC+ on acute medical hospitalizations.
Claims-based quality of care outcomes	
Other quality of care	
Annualized number of low-value services per 1,000 beneficiaries	• Low-value services are medical treatments and procedures that are frequently ordered or prescribed but offer limited benefits to many patients. Examples of low-value services include CT scans for uncomplicated, acute rhinosinusitis and PSA testing for men older than 75 years. Though not an explicit part of CPC+, CPC+ practices nevertheless may be motivated to reduce their patients' use of low-value services (provided by any of their providers) because doing so could help the practices achieve better patient outcomes and reduce the cost of care.

Table 5.C.2. ***(continued)***

[a] We grouped hospitalizations into three categories of severity based on Medicare Severity-Diagnosis Related Group coding: hospitalizations where major complications or comorbidities (MCCs) are present, hospitalizations where complications or comorbidities (CCs) are present, and hospitalizations without a complication or comorbidity present.

CT = computed tomography; PBPM = per beneficiary per month; PSA = prostate-specific antigen; SSP = Medicare Shared Savings Program.

A. Medicare expenditures

In this section, we describe the expenditure outcomes we examined in the impact analysis. First, we present expenditure measures for Medicare Parts A and B; then we discuss Medicare expenditures by service category.

A.1. Medicare expenditures for Part A and Part B services

CMS theorized that changes in care delivery made by CPC+ practices would ultimately result in a reduction in overall Medicare expenditures great enough to offset CMS's enhanced payments. Therefore, we analyzed Medicare expenditures for fee-for-service (FFS) beneficiaries with and without CMS's enhanced payments. All Medicare expenditures exclude third-party and beneficiary liability payments. We provide detailed descriptions for the three Medicare Part A and Part B expenditures measures below. But first we describe the adjustments included in expenditures without enhanced payments and also what counts as enhanced payments.

Medicare expenditures without enhanced payments include Medicare Part A and Part B payments as well as Quality Payment Program (QPP) payments. Starting in 2019 and through 2021, QPP payments include claims-based adjustments for the Merit-based Incentive Payment System (MIPS) that are negative or positive adjustments to physician fees and Critical Access Hospital (CAH) claims and Advanced Alternative Payment Model (APM) incentive payments based on performance two years prior (2017, 2018, and 2019). The MIPS adjustments are included in the payment amount in the 2019, 2020, and 2021 Medicare claims, for performance in 2017, 2018, and 2019, respectively. APM incentive payments are NPI-level payments paid directly to eligible practitioners. We use an NPI-level payment file we received from CMS and a list of NPIs affiliated with each practice. We used random assignment to assign NPIs working at multiple practices to a unique practice and aggregated the NPI level payments to the practice level.[23] For Track 2 practices, CMS also provided alternative payments, in the form of CPCPs, which shifted a portion of the payments practices receive for services from FFS to prospective payments. As these are payments *for services*, they are included in the Medicare expenditure measures without enhanced payments.

Enhanced payments are made *in addition to traditional payments for services and the QPP payments described in the previous paragraph.* As our goal is to estimate impacts for Medicare expenditures for FFS beneficiaries, we do not include enhanced payments from other (non-Medicare) payers in our calculations. Medicare enhanced payments include CMS's CPC+ care management fees (CMFs) for Medicare FFS beneficiaries as well as CMS's payments for rewarding performance. Payments for rewarding performance are: (1) a comprehensiveness supplement for practices participating in Track 2, which is equal to 10 percent of their share of payments (for services) that are made prospectively; (2) prospectively paid and retrospectively reconciled Performance-based Incentive Payments (PBIPs) for

[23] In the third annual report, the proportion of NPIs that worked at multiple practices was 5.2 percent and accounted for 6.5 percent of APM incentive payments.

practices not participating in the Medicare Shared Savings Program (SSP); and (3) shared savings payments to Accountable Care Organizations (ACOs) for practices participating in SSP.

As described below, the three measures of Medicare Part A and Part B expenditures that we include in our impact analysis are: (1) expenditures without enhanced payments; (2) expenditures that include CMFs and the comprehensiveness supplement; and (3) expenditures that include the CMFs, the comprehensiveness supplement, PBIPs, and shared savings payments.

Medicare expenditures for all Part A and Part B services, without enhanced payments, in dollars per beneficiary per month.[24] This measure reflects Medicare expenditures for Part A and Part B covered services during the baseline or intervention period. It includes Medicare payments for inpatient, outpatient, and physician and non-physician services, as well as skilled nursing facilities (SNFs), home health, hospice services, and durable medical equipment (DME) services. Medicare Parts A and B expenditures also include QPP payments and exclude third-party and beneficiary liability payments. The sum of expenditures by service category does not equal the total expenditures for traditional services without enhanced payments, because the total expenditures include lump-sum incentive payments that are not applied at the claim level and instead paid out directly to eligible practitioners who participated in Advanced APMs in 2017, 2018, and 2019.

To obtain the per beneficiary per month (PBPM) amount, we summed Part A and Part B payments for the months a beneficiary was eligible for Medicare FFS during the year and then divided the payments by the number of months the beneficiary was eligible for Medicare FFS. For Track 2 practices, we also included the base CPCPs (but not the 10 percent comprehensiveness supplement). We calculated this PBPM for Track 2 by dividing the total CPCPs to a practice during the reporting period, minus any adjustments or debits (due to retrospective changes in Medicare FFS eligibility of attributed beneficiaries or duplicative billing of services) or recoupments due to early withdrawal from the model, by the total number of Medicare FFS eligible beneficiary-months among beneficiaries assigned to that practice during the period.

Medicare expenditures for all Part A and Part B services, including the CMFs and the comprehensiveness supplement**, in dollars PBPM.** We added the following payments to the expenditures measure (in dollars PBPM):

- The net care management fees (after accounting for debits and recoupments)[25]
- The 10 percent comprehensiveness supplement, for Track 2 practices only

[24] We do not include Part D expenditures, because Medicare makes prospective payments to Part D prescription drug plans that are not directly related to each individual prescription filled by a beneficiary. That is, changes in beneficiaries' prescription use do not affect their PBPM Medicare expenditures.

[25] CMS paid practices in Track 1 and Track 2 average CMFs of $15 and $28, respectively, per month per attributed CPC+ beneficiary in Medicare FFS. These fees were higher than the average fees per month across all intervention years (2017-2021) received of $12 and $23 PBPM for Track 1 and Track 2 practices, respectively, in our analysis sample, because (1) our ITT sample follows beneficiaries even after they are no longer attributed to a CPC+ practice and therefore the practice is no longer receiving CMFs for the Medicare FFS beneficiary, and (2) the list of practitioners and the attribution approach we use for the evaluation are slightly different from those CMS uses for payment. This slight discrepancy between average CMS payments and average payments in our ITT sample applies to PBIPs as well as Track 2 CPCPs. Therefore, all our calculated PBPM payment amounts (for CMFs and PBIPs in both tracks, and for CPCPs in Track 2) for the analysis sample are lower than the CMS-reported numbers for the intervention sample.

Starting in PY 1 (2017), CPC+ practices in both tracks received CMFs from CMS, in addition to usual payments for services, to support their participation in CPC+. CMFs are paid to practices at regular intervals—most commonly at the beginning of each quarter or month—for each patient a payer partner attributes to a practice.

Medicare expenditures for all services, including the CMFs, the comprehensiveness supplement, PBIPs, and SSP payments, in dollars PBPM. We added enhanced payments to the expenditures measure directly above. Specifically, we added the following:

- The final, reconciled PBIP (after recoupments for not meeting quality or utilization targets) for the year received by non-SSP practices
- The shared savings payments earned by their SSP ACO for the SSP practices

For each practice, we divided the CMFs, the 10 percent comprehensiveness supplement, and the PBIPs by the total number of Medicare FFS eligible beneficiary-months in the practice during the reporting period to get the PBPM amounts. There were three steps for adjusting Medicare expenditures for SSP ACO payments. First, we identified the beneficiaries in our sample that were part of an SSP ACO (as determined by the beneficiary level participation data available through MDM). Next, we divided the total shared savings payments earned by their SSP ACO during the reporting period by the total number of Medicare FFS eligible beneficiary-months in that ACO during the period to get a PBPM amount. Lastly, we added this PBPM amount to the average monthly expenditure calculated for these beneficiaries. For example, if an ACO received $500,000 in shared savings and had 50,000 Medicare FFS beneficiary months associated with it for that year (e.g., 5,000 beneficiaries with an average of 10 months of Medicare FFS coverage leading to 50,000 beneficiary months), then we first calculated the PBPM amount of shared savings as $10 PBPM. If only 500 of those beneficiaries in the ACO were also attributed to a CPC+ or comparison practice, then for each of those 500 beneficiaries in our analysis sample, we added $10 PBPM to their claims-based PBPM Medicare expenditures amount for that year.

A.2. Medicare expenditures by service category

In addition to analyzing total expenditures, we also report Medicare expenditures for specific services. We exclude enhanced CPC+ payments when examining measures for each service category. However, MIPS adjustments are included in both Part B expenditures and CAH expenditures that are part of the outpatient expenditures, and CPCPs are included in the Part B expenditures. We create measures for Medicare expenditures stratified by type of Part A or Part B service for the service categories below:

- Inpatient facility expenditures include Part A payments for both acute and non-acute hospitalizations. Short-stay, or acute care hospitalizations and CAH claims, are the most frequent (more than 90 percent of the inpatient claims). Non-acute hospitalizations are primarily at psychiatric or rehabilitation hospitals or units.
- Outpatient facility Part A payments include, but are not limited to, hospital outpatient departments (including emergency rooms), Rural Health Clinics (RHCs) and Federally Qualified Health Centers (FQHCs), renal dialysis facilities, outpatient rehabilitation facilities, comprehensive outpatient rehabilitation facilities, and community mental health centers.

- Part B expenditures for services provided by physicians or non-physicians are expenditures for services provided by professional providers, including physicians, physician assistants (PAs), clinical social workers, nurse practitioners (NPs), and clinical nurse specialists (CNSs). Part B expenditures also include some organizational providers, such as freestanding facilities. Examples of these organizational providers include independent clinical laboratories, ambulance providers, freestanding ambulatory surgical centers, and freestanding radiology centers.
- Home health expenditures include both Part A and Part B expenditures paid to Medicare home health agency providers.
- Skilled nursing facility expenditures include Medicare Part A payments for inpatient stays for nursing care, rehabilitation, and other related health services for patients who need nursing care but do not require hospitalization.
- Hospice expenditures are Part A payments to Medicare certified hospices providers.
- Durable medical equipment expenditures are Part B Medicare payments for Medicare-covered equipment. DME prescribed by a primary care practitioner is covered by Part B, while DME received during a SNF or hospital inpatient stay is paid through Medicare Part A and is not included in these expenditures.

In addition, we created a few specific expenditure categories within these broad service categories above for services, such as acute inpatient, inpatient rehabilitation facilities, outpatient emergency department, and ambulatory visits with primary care practitioners and specialists. We describe these more granular expenditure outcomes below.

We created two subsets of outcomes of inpatient expenditures: acute hospitalization expenditures (total and also broken down by surgical or medical) and inpatient rehabilitation facility expenditures. (As described above, the total inpatient expenditures also include other types of expenditures such as psychiatric hospitalizations). These expenditure measures are described below.

Acute hospitalization expenditures. The first is short-stay acute inpatient/CAH expenditures. We categorized an inpatient stay as a short-stay acute inpatient hospital stay when the third through sixth digits of the provider number are equal to 0001 through 0899. If the third and fourth digits of the provider number are equal to 13, then it is a CAH stay.

We further divided short-stay acute expenditures based on whether they were expenditures for surgical or medical (that is, non-surgical admissions), using details on the list of MS-DRGs from Table 5 on the IPPS Final Rule page for each year from 2016 to 2022. Specifically, we used the variable "MS-DRG Type," which indicates whether the admission was a surgical or a medical MS-DRG:

- **Medicare expenditures for acute surgical hospitalizations.** This measure is the average monthly expenditure for acute hospitalizations with MS-DRG type "SURG" in the year, expressed as per beneficiary per month.
- **Medicare expenditures for acute medical (i.e., non-surgical) hospitalizations.** This measure is the average monthly expenditure for acute hospitalizations with MS-DRG type "MED" in the year, expressed as per beneficiary per month.

Inpatient rehabilitation facility expenditures. The second subset of inpatient expenditures is Medicare payments for inpatient rehabilitation facilities (IRFs). IRF claims are identified using the provider number values 3025 through 3099 in the third through sixth digit or if there is a value of R or T in the third position. Note that IRF expenditures are a subset of the non-acute hospitalization component of total inpatient expenditures. The remaining expenditures for other non-acute facilities are not reported separately.

Medicare expenditures for post-acute care (PAC). We created two measures that identify PAC expenditures incurred after an acute hospitalization using inpatient, IRF, SNF, and home health care claims. One is a standalone PAC summary of expenditures and the second includes the index admission expenditures associated with the PAC episode. We defined post-acute care expenditures as the expenditures associated with care delivered during a sequence of post-acute stays (a "PAC episode") for which each stay is separated from previous stays by no more than seven days.[26] Once all stays for a PAC episode were identified, we summed the total Medicare payments for each PAC stay to identify the total PAC episode expenditure amount.[27]

A PAC episode begins with the discharge from an index inpatient hospitalization. The same methodology used to identify index hospitalizations for the unplanned readmissions measure is also applied here (see Section E). Some acute or CAH stays are excluded because they meet certain criteria, such as hospitalizations for rehabilitation, the medical treatment of cancer, or the patient left against medical advice. Those that do not meet exclusion criteria are retained and considered index inpatient hospitalizations for a PAC episode when they meet one of the following criteria:

- Within seven days of the index hospitalization discharge, a beneficiary is admitted to an IRF or long-term care hospital (LTCH), or
- If the index hospitalization is no less than three days long, a beneficiary:
  - Is admitted to an SNF no more than 30 days after the discharge date,[28] or
  - Receives home health care services no more than 14 days after the discharge date.

The identification of PAC services in the Medicare claims and the definition of stays within each of these PAC services are provided in Table 5.C.3.

[26] Our definition of a PAC episode is based on the definition used by Urban Institute in its report to the Medicare Payment Advisory Commission. https://www.medpac.gov/wp-content/uploads/import_data/scrape_files/docs/default-source/contractor-reports/sept2018_pac_sequence_of_care_w_cov_contractor_sec.pdf.

[27] Acute inpatient stays in the middle of PAC episodes were included in identifying a sequence of stays with no more than seven days separating one stay from the next, but expenditures associated with those intervening inpatient stays were not included in PAC expenditures.

[28] SNF stays that met the three-day rule waiver (that is, stays with a claim demonstration code of 68) are not included in our definition of a PAC episode when these stays cannot be attributed to an index inpatient hospitalization. In 2015 and 2016, there were 1,427 SNF stays with the demonstration code of 68 that did not have an index inpatient hospitalization associated with them (7.2 percent), affecting 661 out of 10 million beneficiaries and less than 0.01 percent of the more than 16 million PAC episodes identified.

Table 5.C.3. Definitions of PAC claims and stays

Claims type	Claims included	Stay definition[a]
IRF	Provider number values from the inpatient file: • 3025 through 3099 in the third through sixth digit; or • R or T in the third position	Separate IRF stays were created if there is a hospitalization or PAC stay of greater than three days or the beneficiary returned to a different IRF facility
Home health	All claims from home health file	Sixty consecutive days are considered a single home health stay even if an intervening hospital or institutional stay occurs within the 60-day episode
LTCH	Provider number values from the inpatient file: • 20 through 22 in the third and fourth digit	Separate stays are created if the patient returned to the same LTCH following a stay in: • An acute hospital for at least 10 days, • An IRF for at least 28 days, or • An SNF for at least 46 days
SNF	All claims from the SNF file and the following provider number values from the inpatient file: • 5000 through 6499 in the third through sixth digit; or • U, W, Y, or Z in the third position	Any discharges and later admissions to an SNF resulted in a separate SNF stay, regardless of any care the beneficiary received between the two SNF claims

[a] A stay is specific to each PAC type.

IRF = inpatient rehabilitation facility; LTCH = long-term care hospital; PAC = post-acute care; SNF = skilled nursing facility.

Outpatient ED (including observation stays) expenditures. We created an outpatient facility and professional expenditures measure for emergency department (ED) claims that is a subset of total hospital outpatient department expenditures. To identify outpatient ED visits for this expenditure measure, we use the approach described in the service utilization section below, with one exception: expenditures are not restricted to one ED stay per day, to ensure we include all expenditures associated with these services. We used a two-step process to identify professional expenditures associated with outpatient facility ED claims. First, we identified professional claims with a place of service code equal to 2, which indicates ED or an evaluation and management service provided in the ED (CPT code equal to 99281-99285) or during an observation stay (CPT code equal to 99217-99220 or 99224-99226). Next, we linked these professional claims to outpatient facility ED claims and retained professional claims with dates of service overlapping or one day before or after the dates of service in an outpatient facility ED claim for the same beneficiary.

Medicare expenditures for ambulatory visits. We also identified expenditures for ambulatory visits using carrier claims and FQHC, RHC, and CAH claims from the outpatient file. Note that visits associated with the carrier file do not include potential facility fees. We created two categories of ambulatory visit expenditures: (1) ambulatory visits with primary care practitioners and (2) ambulatory visits with specialists. For ambulatory services provided by primary care practitioners, we further calculated expenditures for services provided by primary care practitioners at the beneficiary's assigned practice versus at other practices. Finally, we also examined the proportion of ambulatory visit expenditures with primary care practitioners that are non-face-to-face (includes telephone, online via a secure platform, or other audio or video connection). See Section B.4 for more details on non-face-to-face ambulatory physician visits.

Laboratory and imaging expenditures. We created these two outcomes to capture expenditures for diagnostic procedures using carrier and outpatient claims. For laboratory services, we selected procedures that were in the 80000–89999 range (clinical laboratory tests or examinations) or were assigned a BETOS

code starting with "T1" (lab tests), or an RBCS[29] code starting with "TA" (anatomic pathology), "TL" (general laboratory), or "TM" (molecular testing). For imaging services, we selected procedures that were in the 70000–79999 range (radiology and pathology) or were assigned a BETOS or RBCS code starting with "I."

B. Service use

We evaluated impacts on a range of service use outcomes for Medicare FFS beneficiaries, so that CMS might consider the patterns of effects across these domains along with any observed impacts on Medicare expenditures without and with CMS's enhanced payments. These selected measures of Medicare service use include the number of acute hospitalizations, ED visits, urgent care center (UCC) visits, ambulatory visits, and other service use, such as 30-day unplanned readmissions.

B.1. Acute hospitalizations

Number of hospitalizations at short-stay acute hospitals and CAHs per 1,000 beneficiaries per year. This measure is the annualized hospitalization rate per 1,000 beneficiaries of all short-stay acute hospital and CAH admissions. Transfers between acute/CAH facilities are counted as a single admission. Multiple claims for acute admissions from traditional acute care hospitals and CAHs that represent transfers between hospitals are combined into a single record, so that they count as one admission.

As with expenditures for acute hospitalizations, we divided the number of acute hospitalizations into two measures by whether they were for surgical or medical hospitalizations, using the same variable MS-DRG type from the IPPS Final Rule, as described above:

- **Number of acute surgical hospitalizations.** For this measure, we counted the number of hospitalizations in the year that had an MS-DRG type "SURG." We expressed it as the annualized rate per 1,000 beneficiaries.
- **Number of acute medical hospitalizations.** For this measure, we counted the number of hospitalizations in the year that had an MS-DRG type "MED." We expressed it as the annualized rate per 1,000 beneficiaries.

We then subdivided the acute surgical and acute medical hospitalization measures into three severity groupings using the *MS-DRG Description variable,* in Table 5 of the IPPS Final Rule page, which provides a short description of the MS-DRG. For example, the description of MS-DRG 177 is "Respiratory infections and inflammations with a major complication or comorbidity." Many descriptions indicate whether the MS-DRG has a complication or comorbidity (CC) or a major complication or comorbidity (MCC), as in the example above. CMS makes this distinction because diagnoses with MCCs require more resources to treat than less-major comorbid conditions do. Using the combination of MS-DRG type and MS-DRG Description, we created six additional outcomes:

- **Number of acute surgical hospitalizations with an MCC.** For this measure, we counted the number of hospitalizations in the year that had an MS-DRG with MS-DRG type "SURG" and an MS-DRG description of an MCC. We expressed it as the annualized rate per 1,000 beneficiaries.

[29] In 2021, CMS developed the Restructured BETOS Classification System (Templeman et al. 2020) to refine and update the structure of the existing BETOS service categorizations.

- **Number of acute surgical hospitalizations with a CC.** For this measure, we counted the number of hospitalizations in the year that had an MS-DRG with MS-DRG type "SURG" and an MS-DRG description of a CC. We expressed it as the annualized rate per 1,000 beneficiaries.
- **Number of acute surgical hospitalizations without an MCC/CC.** For this measure, we counted the number of hospitalizations in the year that had an MS-DRG with MS-DRG type "SURG" and an MS-DRG description that the admission did not have an MCC/CC. We expressed it as the annualized rate per 1,000 beneficiaries.
- **Number of acute medical hospitalizations with an MCC.** For this measure, we counted the number of hospitalizations in the year that had an MS-DRG with MS-DRG type "MED" and an MS-DRG description of an MCC. We expressed it as the annualized rate per 1,000 beneficiaries.
- **Number of acute medical hospitalizations with a CC.** For this measure, we counted the number of hospitalizations in the year that had an MS-DRG with MS-DRG type "MED" and an MS-DRG description of a CC. We expressed it as the annualized rate per 1,000 beneficiaries.
- **Number of acute medical hospitalizations without an MCC/CC.** For this measure, we counted the number of hospitalizations in the year that had an MS-DRG with MS-DRG type "MED" and an MS-DRG description that the admission did not have an MCC/CC. We expressed it as the annualized rate per 1,000 beneficiaries.

B.2. ED visits

Number of ED visits per 1,000 beneficiaries per year. We created an overall ED visit measure that combines ED visits leading to a hospitalization with outpatient ED visits (and observation stays). Note that an observation stay, by definition, does not always lead to an inpatient admission. In addition, we reported the outpatient ED visits separately. We describe the methodology for identifying the two components of this measure below.

ED visits that lead to a hospitalization are identified in the inpatient file and include acute, critical access, or psychiatric hospital stays that have a claim with a revenue center line item equal to 045X or 0981 (emergency room care) or 0762 (treatment or observation room). These visits are not shown separately.

Outpatient ED visits are identified in the outpatient department 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 Health Care Common Procedure Coding System (HCPCS) code of G0378 (hospital observation services per hour). If the procedure code on the line item of the ED claim was equal to 70000 through 79999 or 80000 through 89999, we excluded it; this exclusion was intended to exclude claims in which only radiological or pathology/laboratory services were provided. We then capped the number of ED visits to one per day.

Outpatient ED visits, including observation stays, per 1,000 beneficiaries per year. This measure is the annualized number of emergency room visits and observation stays (combined to create ED visits) that do not lead to a hospitalization, per 1,000 beneficiaries.

Because the COVID-19 pandemic drove shifts in the distribution of primary diagnosis codes on ED visit claims in 2020 and 2021 relative to prior years, to test the sensitivity of this measure to these shifts we developed an additional version of this measure that excluded ED visits with a primary diagnosis of

COVID-19, exposure to COVID-19, other viral or biological agent, and respiratory illness (see Table 5.C.4 for specific diagnosis codes excluded).

Table 5.C.4. Diagnosis codes associated with the COVID-19 pandemic

Diagnosis code	Description
COVID-19 diagnosis or potential exposure to COVID-19	
B9729	Other coronavirus
U071	2019 novel coronavirus
J12.82	Pneumonia due to coronavirus disease
Z20822	Contact with and (suspected) exposure to COVID-19, SARS-COV-2, or other novel coronavirus
Z20828	Contact with and (suspected) exposure to other viral communicable diseases
Z03818	Exposure to biological agent
Z1152	Encounter for screening for other viral disease
Z01818	Encounter for other preprocedural exam
Z01812	Encounter for preprocedural lab
Z1159	Encounter for observation for suspected exposure to other biological agents ruled out
Respiratory illness codes	
J1289	Viral pneumonia
J208	Bronchitis – acute
J40	Bronchitis – unspecified
J988	Lower respiratory infection – specified
J22	Lower respiratory infection – unspecified
J80	Acute respiratory distress syndrome
J069	Acute upper respiratory infection, unspecified
R05	Cough
J209	Acute bronchitis
R059	Cough, unspecified

Primary care substitutable ED visits and potentially primary care preventable outpatient ED visits—each calculated per 1,000 beneficiaries per year. These measures are subsets of the outpatient ED visits identified above. The construction of these measures aligns with the New York University Emergency Department Algorithm (NYU EDA), the measure most commonly used to identify primary care treatable ED visits. To this algorithm, we applied the "patch" developed by Johnston et al. (2017) that updates the algorithm with ICD-9 and ICD-10 codes added since 2001. This algorithm assigns all ED visits identified for the outpatient ED visit measure above the probability of the visit being in each of the following categories: (1) nonemergent; (2) emergent but treatable in a primary care setting; (3) emergent/ED care required but preventable or avoidable if appropriate ambulatory care had been received; and (4) emergent/ED care required and not preventable or avoidable. If there are multiple ED claims with the same from date, we keep only the first claim to appear in the file.

- The probability of a visit being primary care substitutable is calculated as the sum of the probabilities that the visit is nonemergent or emergent but treatable in a primary care setting (NYU Categories 1 and 2).
- The probability of a visit being potentially primary care preventable is calculated as the sum of the probabilities for the categories in which the visit is emergent and ED care is required (Categories 3 and 4).

We summed these probabilities across all ED visits to estimate the total number of primary care substitutable ED visits and the total number of potentially primary care preventable ED visits.

We created two versions of this measure, overall and excluding services primarily for a COVID-19 diagnosis or a respiratory condition potentially caused by COVID-19. (See Table 5.C.4 for specific diagnosis codes excluded.)

B.3. Urgent care center visits

Total urgent care center (UCC) visits per 1,000 beneficiaries per year. This measure includes UCC visits identified in the carrier claims file based on a place of service equal to 20 and outpatient hospital file services with a revenue code of 516 or 526. If there are multiple UCC visits with the same initial date of service, we counted only the first UCC claim to appear in the file. We created two versions of this measure, overall and excluding services primarily for a COVID-19 diagnosis or a respiratory condition potentially caused by COVID-19. (See Table 5.C.4 for specific diagnosis codes excluded.)

Primary care substitutable UCC visits per 1,000 beneficiaries per year. Like the parallel ED visit measure described above, the construction of this measure aligns with the NYU EDA. To the NYU EDA, we applied the "patch" developed by Johnston et al. (2017). We used this algorithm to assign all UCC visits identified for the total UCC visit count measure above the probability of the visit being in each of the following categories: (1) nonemergent; (2) emergent but treatable in a primary care setting; (3) emergent/ED care required but preventable or avoidable if appropriate ambulatory care had been received; and (4) emergent/ED care required and not preventable or avoidable. If there are multiple UCC claims with the same from date, we keep only the first claim to appear in the file. We calculated the probability of a UCC visit being primary care substitutable by summing the probabilities that the visit is in the nonemergent or emergent but treatable in a primary care setting categories. We summed these probabilities across all UCC visits to estimate the total number of primary care substitutable UCC visits. We created two versions of this measure, overall and excluding services primarily for a COVID-19 diagnosis or a respiratory condition potentially caused by COVID-19. (See Table 5.C.4 for specific diagnosis codes excluded.)

B.4. Ambulatory visits, including visits to FQHCs, RHCs, and CAHs

We created two measures of the number of ambulatory visits: annualized visits per 1,000 beneficiaries to (1) primary care practitioners and (2) specialists. Specialties were grouped into primary care practitioners and specialists as defined by Healthcare Provider Taxonomy Codes reported in the National Plan and Provider Enumeration System (NPPES) (taxonomy codes are listed in Table 5.C.5 for primary care practitioners and in Table 5.C.6 for specialists). Multiple claims with the same practitioner on the same day are counted as one visit, and multiple claims with different practitioners on the same day are counted as separate visits. We discuss the criteria for identifying ambulatory visits and updates to the methodology since our second annual report below:

- To identify a practitioners' specialty, we use only the primary taxonomy code from the NPPES, rather than both the primary and secondary taxonomy codes (a change implemented in the second annual report).
- In the third annual report, we identified new specialties for primary care practitioners and specialists to ensure consistency across measures that use specialty designations. The specialty designations are now the same across the measures of ambulatory visits and continuity/fragmentation of care (see

Appendix 5.C in Orzol et al. [2021] for details). There were no changes to specialty designations for this report.

- In the fourth annual report, we expanded our definition of ambulatory visits to align with the narrow definition of primary care services that others have used to measure primary care spending in both the Medicare and the commercially insured populations (Bailit et al. 2017; Reid et al. 2019; Kempski and Greiner 2020). This definition includes procedure codes for professional claims, including evaluation and management visits, preventive visits, care transition or coordination services, and in-office preventive services, screening, and counseling. Table 5.C.7 provides a complete list of visits for office-based evaluation and management, nursing home and home care, care management services (including behavioral health), health and behavior assessments, psychotherapy, and other services mentioned above—as defined by HCPCS/Current Procedural Terminology (CPT) and revenue center codes. Table 5.C.8 explains the codes.
- Add-on services are counted in the expenditures but not in utilization measures as a separate service (creating a more precise count of actual ambulatory visits). For example, CPT code 99354 is for prolonged physician services in an office or outpatient setting billed on the same day as the companion evaluation and management codes (e.g., office or other outpatient E&M visits). See the Ambulatory Visit Indicator column in Table 5.C.8 for the complete list of visits identified as "add-on" services.
- Certain services qualify only if they have a non-inpatient place of service to limit to services in ambulatory settings only (primarily, newly added behavioral health services). Table 5.C.8 identifies procedure codes subject to these additional criteria in the Place of Service Indicator column.
- Ambulatory visits on the outpatient file are included only if they were provided at an FQHC, RHC, or CAH, to avoid double-counting services that would appear in the physician bills on the carrier file.
- The CPT Editorial Panel instituted several procedure code updates during our analytic time period. Therefore, we updated our specifications to reflect codes as they were added, deleted, or replaced. We included new procedure codes as they were implemented or updated them when they were replaced. These changes are tracked in Table 5.C.9.

Number of ambulatory visits to primary care practitioners (including visits to FQHCs, RHCs, and CAHs) per 1,000 beneficiaries per year. This measure is the number of annualized ambulatory visits per 1,000 beneficiaries to primary care practitioners, including physicians, NPs, CNSs, and PAs. Table 5.C.5 lists primary care-specific taxonomy codes. Codes for ambulatory visits are listed in Table 5.C.7 and explained in Table 5.C.8.

Number of ambulatory visits to specialists (including visits to FQHCs, RHCs, and CAHs) per 1,000 beneficiaries per year. This measure is the number of annualized ambulatory visits per 1,000 beneficiaries to specialists, including surgeons, psychiatrists, and emergency medicine practitioners. Table 5.C.6 lists specialty taxonomy codes. We exclude non-specialist taxonomies, such as laboratories, ambulance, chiropractor, and physical therapy. To identify the number of specialist ambulatory visits, we use the same criteria we use to identify ambulatory visits to primary care practitioners. Codes for ambulatory visits are listed in Table 5.C.7 and explained in Table 5.C.8.

Proportion of ambulatory visits to primary care practitioners that were non-face-to-face (including visits to FQHCs, RHCs, and CAHs) per 1,000 beneficiaries per year. Given the increase in telehealth use during the COVID-19 pandemic, we examined the proportion of ambulatory visits that were non-face-to-face in PYs 4 and 5. We identified a subset of ambulatory visits as non-face-to-face using three selection criteria. All remaining ambulatory visits are considered face-to-face encounters. Non-face-to-face ambulatory visits are:

1. Ambulatory visit procedure codes such as telephone and online E&M; telephone and online assessment and management; chronic care remote patient monitoring; and virtual check-ins. These codes are in green shaded rows for easy identification in Table 5.C.8.
2. Ambulatory visits with a modifier value of 95, GT, GQ, or G0 indicating a telehealth visit.
3. Ambulatory visits identified on the carrier file that have the place of service equal to 02 (telehealth).

Table 5.C.5. Primary care taxonomy codes

Medicare practitioner-type description	Practitioner taxonomy code	Practitioner taxonomy description
Physician/Family Practice	207Q00000X	Physicians/Family Medicine
	207QA0000X	Physicians/Family Medicine, Adolescent Medicine**
	207QA0505X	Physicians/Family Medicine, Adult Medicine
	207QG0300X	Physicians/Family Medicine, Geriatric Medicine
Physician/Internal Medicine	207R00000X	Physicians/Internal Medicine
	207RA0000X	Physicians/Internal Medicine, Adolescent Medicine**
	207RG0300X	Physicians/Internal Medicine, Geriatric Medicine
Physician/Pediatrics[a]	208000000X	Physicians/Pediatrics**
	2080A0000X	Physicians/Pediatrics, Adolescent Medicine**
	2080P0006X	Physicians/Pediatrics, Developmental/Behavioral Pediatrics***
	2080P0008X	Physicians/Pediatrics, Neurodevelopmental Disabilities***
	2083B0002X	Physicians/Pediatrics, Preventative Medicine***
Nurse Practitioner	363L00000X	Nurse Practitioner
	363LA2100X	Nurse Practitioner, Acute Care
	363LA2200X	Nurse Practitioner, Adult Health
	363LC1500X	Nurse Practitioner, Community Health
	363LF0000X	Nurse Practitioner, Family
	363LG0600X	Nurse Practitioner, Gerontology
	363LP0200X	Nurse Practitioner, Pediatrics**
	363LP2300X	Nurse Practitioner, Primary Care
	363LW0102X	Nurse Practitioner, Women's Health
Certified Clinical Nurse Specialist	364S00000X	Clinical Nurse Specialist
	364SA2100X	Clinical Nurse Specialist, Acute Care
	364SA2200X	Clinical Nurse Specialist, Adult Health
	364SC1501X	Clinical Nurse Specialist, Community Health/Public Health
	364SC2300X	Clinical Nurse Specialist, Chronic Care
	364SF0001X	Clinical Nurse Specialist, Family Health
	364SG0600X	Clinical Nurse Specialist, Gerontology
	364SH1100X	Clinical Nurse Specialist, Holistic
	364SP0200X	Clinical Nurse Specialist, Pediatrics**
	364SW0102X	Clinical Nurse Specialist, Women's Health
Physician Assistant	363A00000X	Physician Assistant
	363AM0700X	Physician Assistant, Medical
Physician/Undefined Physician Type	208D00000X	General Practice
	2083P0901X	General Practice, Public Health & General Preventive Medicine***
Federally Qualified Health Center	261QF0400X	Ambulatory Health Care Facilities/FQHC
Rural Health Clinic	261QR1300X	Ambulatory Health Care Facilities/Clinic Center, Rural Health

Source: Centers for Medicare & Medicaid Services. "Crosswalk Medicare Provider/Supplier to Healthcare Provider Taxonomy." Baltimore, MD: CMS. Available at https://data.cms.gov/Medicare-Enrollment/CROSSWALK-MEDICARE-PROVIDER-SUPPLIER-to-HEALTHCARE/j75i-rw8y. Accessed January 4, 2022.

Notes: Descriptions annotated with two asterisks (**) are categories added since our first annual report; three asterisks (***) indicate categories that have been added since our second annual report. To ensure consistency across measures that use specialty designations, we identified new specialties for primary care practitioners in our second and third annual reports. The specialty designations remain the same across the measures of ambulatory visits and continuity/fragmentation of care measures. Taxonomy code 207QH0002X (Hospice and Palliative Medicine) was removed and added to specialist care in the second annual report.

[a] This Physician/Pediatrics specialty is more relevant for analyses of the Medicaid population, but it will also capture some beneficiaries in the Medicare population.

Table 5.C.6. Specialist care taxonomy codes

Medicare practitioner-type description	Practitioner taxonomy code	Practitioner taxonomy description
Surgery	208600000X	Physicians/Surgery
	2086S0120X	Physicians/Surgery/Pediatric Surgery
	2086S0122X	Physicians/Surgery/Plastic and Reconstructive Surgery
	2086S0105X	Physicians/Surgery/Surgery of the Hand
	2086S0102X	Physicians/Surgery/Surgical Critical Care
	2086X0206X	Physicians/Surgery/Surgical Oncology
	2086S0127X	Physicians/Surgery/Trauma Surgery
	2086S0129X	Physicians/Surgery/Vascular Surgery
	208G00000X	Physicians/Thoracic
	204F00000X	Physicians/Transplant Surgery
	208C00000X	Physicians/Colon & Rectal Surgery
	207T00000X	Physicians/Neurological Surgery
	204E00000X	Physicians/Oral & Maxillofacial Surgery
	207X00000X	Physicians/Orthopedic Surgery
	207XS0114X	Physicians/Orthopedic Surgery/Adult Reconstructive Orthopedic Surgery
	207XX0004X	Physicians/Orthopedic Surgery/Foot and Ankle Surgery
	207XS0106X	Physicians/Orthopedic Surgery/Hand Surgery
	207XS0117X	Physicians/Orthopedic Surgery/Orthopedic Surgery of the Spine
	207XX0801X	Physicians/Orthopedic Surgery/Orthopedic Trauma
	207XP3100X	Physicians/Orthopedic Surgery/Pediatric Orthopedic Surgery
	207XX0005X	Physicians/Orthopedic Surgery/Sports Medicine
	208200000X	Physicians/Plastic Surgery
	2082S0099X	Physicians/Plastic Surgery/Plastic Surgery Within the Head & Neck
	2082S0105X	Physicians/Plastic Surgery/Surgery of the Hand
	2086H0002X	Physicians/Surgery/Hospice and Palliative Medicine***
Allergy/Immunology/ Otolaryngology	207K00000X	Physicians/Allergy and Immunology
	207KA0200X	Physicians/Allergy and Immunology/Allergy
	207KI0005X	Physician/Allergy and Immunology/Allergist***
	207Y00000X	Physicians/Otolaryngology
	207YS0123X	Physicians/Otolaryngology/Facial Plastic Surgery
	207YX0602X	Physicians/Otolaryngology/Otolaryngic Allergy
	207YX0905X	Physicians/Otolaryngology/Otolaryngology/Facial Plastic Surgery
	207YX0901X	Physicians/Otolaryngology/Otology &Neurotology
	207YP0228X	Physicians/Otolaryngology/Pediatric Otolaryngology
	207YX0007X	Physicians/Otolaryngology/Plastic Surgery within the Head & Neck
	207YS0012X	Physicians/Otolaryngology/Sleep Medicine***
Anesthesiology	207L00000X	Physicians/Anesthesiology
	207LC0200X	Physicians/Anesthesiology/Critical Care Medicine
	207LP3000X	Physicians/Anesthesiology/Pediatric Anesthesiology
	207RC0000X	Physicians/Internal Medicine, Cardiovascular Disease
	207LA0401X	Physician/Anesthesiology, Addiction Medicine***
	207LH0002X	Physician/Anesthesiology, Hospice and Palliative Medicine***
	207LP2900X	Physician/Anesthesiology, Pain Medicine***

Table 5.C.6. ***(continued)***

Medicare practitioner-type description	Practitioner taxonomy code	Practitioner taxonomy description
Dermatology	207N00000X	Physicians/Dermatology
	207NI0002X	Physicians/Dermatology, Clinical & Laboratory Dermatological Immunology
	207ND0101X	Physicians/Dermatology, MOHS-Micrographic Surgery
	207ND0900X	Physicians/Dermatology, Derma pathology
	207NP0225X	Physicians/Dermatology, Pediatric Dermatology
	207NS0135X	Allopathic &Osteopathic Physicians/Dermatology, Procedural Dermatology
Obstetrics & Gynecology	207V00000X	Physicians/Obstetrics & Gynecology
	207VB0002X	Physicians/Obstetrics & Gynecology, Bariatric Medicine
	207VC0200X	Physicians/Obstetrics & Gynecology, Critical Care Medicine
	207VF0040X	Physicians/Obstetrics & Gynecology, Female Pelvic Medicine and Reconstructive Surgery
	207VX0201X	Physicians/Obstetrics & Gynecology, Gynecologic Oncology
	207VG0400X	Physicians/Obstetrics & Gynecology, Gynecology
	207VM0101X	Physicians/Obstetrics & Gynecology, Maternal & Fetal Medicine
	207VX0000X	Physicians/Obstetrics & Gynecology, Obstetrics
	207VE0102X	Physicians/Obstetrics & Gynecology, Reproductive Endocrinology
	207VH0002X	Physicians/Obstetrics & Gynecology, Hospice and Palliative Medicine***
Ophthalmology	207W00000X	Physicians/Ophthalmology
	207WX0009X	Physicians/Ophthalmology, Glaucoma Specialist
	207WX0107X	Physicians/Ophthalmology, Retina Specialist
	207WX0108X	Physicians/Ophthalmology, Uveitis and Ocular Inflammatory Disease
	207WX0109X	Physicians/Ophthalmology/Neuro-ophthalmology
	207WX0110X	Physicians/Ophthalmology/Pediatric Ophthalmology and Strabismus Specialist
	207WX0120X	Physicians/Ophthalmology, Cornea and External Diseases Specialist
	207WX0200X	Physicians/Ophthalmic Plastic and Reconstructive Surgery
	1223S0112X	Physicians/Ophthalmology, Dental Providers/Dentist, Oral & Maxillofacial Surgery
Pathology	207ZP0101X	Physicians/Pathology, Anatomic Pathology
	207ZP0102X	Physicians/Pathology, Anatomic Pathology & Clinical Pathology
	207ZP0104X	Physicians/Pathology, Chemical Pathology
	207ZC0006X	Physicians/Pathology, Clinical Pathology
	207ZP0105X	Physicians/Pathology, Clinical Pathology/Laboratory Medicine
	207ZC0500X	Physicians/Pathology, Cytopathology
	207ZD0900X	Physicians/Pathology, Derma pathology
	207ZF0201X	Physicians/Pathology, Forensic Pathology
	207ZH0000X	Physicians/Pathology, Hematology
	207ZI0100X	Physicians/Pathology, Immunopathology
	207ZM0300X	Physicians/Pathology, Medical Microbiology
	207ZP0007X	Physicians/Pathology, Molecular Genetic Pathology
	207ZN0500X	Physicians/Pathology, Neuropathology
	207ZP0213X	Physicians/Pathology, Pediatric Pathology
Physical Medicine & Rehabilitation	208100000X	Physicians/Physical Medicine & Rehabilitation
	2081H0002X	Physicians/Physical Medicine & Rehabilitation, Hospice and Palliative Medicine
	2081N0008X	Physicians/Physical Medicine & Rehabilitation, Neuromuscular Medicine
	2081P2900X	Physicians/Physical Medicine & Rehabilitation, Pain Medicine
	2081P0010X	Physicians/Physical Medicine & Rehabilitation, Pediatric Rehabilitation Medicine
	2081P0004X	Physicians/Physical Medicine & Rehabilitation, Spinal Cord Injury Medicine
	2081S0010X	Physicians/Physical Medicine & Rehabilitation, Sports Medicine

Table 5.C.6. ***(continued)***

Medicare practitioner-type description	Practitioner taxonomy code	Practitioner taxonomy description
	2081P0301X	Physicians/Physical Medicine & Rehabilitation, Brain Injury
Urology	208800000X	Physicians/Urology
	2088P0231X	Physicians/Urology, Pediatric Urology
	2088F0040X	Female Pelvic Medicine & Reconstructive Surgery
Internal Medicine	207RN0300X	Physicians/Internal Medicine, Nephrology
	207RP1001X	Physicians/Internal Medicine, Pulmonary Disease
	207RI0200X	Physicians/Internal Medicine, Infectious Disease
	207RE0101X	Physicians/Internal Medicine, Endocrinology, Diabetes & Metabolism
	207RR0500X	Physicians/Internal Medicine, Rheumatology
	207RC0200X	Physicians/Internal Medicine, Critical Care Medicine
	207RH0000X	Physicians/Internal Medicine, Hematology
	207RH0003X	Physicians/Internal Medicine, Hematology & Oncology
	207RX0202X	Physicians/Internal Medicine, Medical Oncology
	207RA0201X	Physicians/Internal Medicine, Allergy & Immunology***
	207RA0401X	Physicians/Internal Medicine, Addiction Medicine***
	207RB0002X	Physicians/Internal Medicine, Bariatric Medicine***
	207RC0001X	Physicians/Internal Medicine, Clinical Cardiac Electrophysiology***
	207RG0100X	Physicians/Internal Medicine, Gastroenterology***
	207RH0002X	Physicians/Internal Medicine, Hospice and Palliative Medicine***
	207RH0005X	Physicians/Internal Medicine, Hypertension Specialist***
	207RI0001X	Physicians/Internal Medicine, Clinical & Laboratory Immunology***
	207RI0008X	Physicians/Internal Medicine, Hepatology***
	207RI0011X	Physicians/Internal Medicine, Interventional Cardiology***
	207RM1200X	Physicians/Internal Medicine, Magnetic Resonance Imaging (MRI)***
	207RS0010X	Physicians/Internal Medicine, Sports Medicine***
	207RS0012X	Physicians/Internal Medicine, Sleep Medicine***
	207RT0003X	Physicians/Internal Medicine, Transplant Hepatology***
Eye & Vision	152W00000X	Eye and Vision Service Providers/Optometrist
	152WC0802X	Eye and Vision Service Providers/Optometrist, Corneal and Contact Management
	152WL0500X	Eye and Vision Service Providers/Optometrist, Low Vision Rehabilitation
	152WX0102X	Eye and Vision Service Providers/Optometrist, Occupational Vision
	152WP0200X	Eye and Vision Service Providers/Optometrist, Pediatrics
	152WS0006X	Eye and Vision Service Providers/Optometrist, Sports Vision
	152WV0400X	Eye and Vision Service Providers/Optometrist, Vision Therapy
Podiatric Medicine	213E00000X	Podiatric Medicine & Surgery Service Providers/Podiatrist
	213ES0103X	Podiatric Medicine & Surgery Service Providers/Podiatrist, Foot & Ankle Surgery
	213ES0131X	Podiatric Medicine & Surgery Service Providers/Podiatrist, Foot Surgery
	213EG0000X	Podiatric Medicine & Surgery Service Providers/Podiatrist, General Practice
	213EP1101X	Podiatric Medicine & Surgery Service Providers/Podiatrist, Primary Podiatric Medicine
	213EP0504X	Podiatric Medicine & Surgery Service Providers/Podiatrist, Public Medicine
	213ER0200X	Podiatric Medicine & Surgery Service Providers/Podiatrist, Radiology
	213ES0000X	Podiatric Medicine & Surgery Service Providers/Podiatrist, Sports Medicine
Psychiatry & Neurology	2084A0401X	Physicians/Psychiatry & Neurology
	2084A2900X	Physicians/Psychiatry & Neurology/Neurocritical Care
	2084P0802X	Physicians/Psychiatry & Neurology, Addiction Psychiatry
	2084B0002X	Physicians/Psychiatry & Neurology, Bariatric Medicine
	2084P0804X	Physicians/Psychiatry & Neurology, Child & Adolescent Psychiatry
	2084N0600X	Physicians/Psychiatry & Neurology, Clinical Neurophysiology

Table 5.C.6. ***(continued)***

Medicare practitioner-type description	Practitioner taxonomy code	Practitioner taxonomy description
	2084D0003X	Physicians/Psychiatry & Neurology, Diagnostic Neuroimaging
	2084F0202X	Physicians/Psychiatry & Neurology, Forensic Psychiatry
	2084P0805X	Physicians/Psychiatry & Neurology, Geriatric Psychiatry
	2084H0002X	Physicians/Psychiatry & Neurology, Hospice & Palliative Medicine
	2084P0005X	Physicians/Psychiatry & Neurology, Neurodevelopmental Disabilities
	2084N0400X	Physicians/Psychiatry & Neurology, Neurology
	2084N0402X	Physicians/Psychiatry & Neurology, Neurology with Special Qualifications in Child Neurology
	2084N0008X	Physicians/Psychiatry & Neurology, Neuromuscular Medicine
	2084P0301X	Psychiatry & Neurology/Respiratory, Developmental, Rehabilitative and Restorative Service, Brain Injury Medicine
	2084P2900X	Physicians/Psychiatry & Neurology, Pain Medicine
	2084P0800X	Physicians/Psychiatry & Neurology, Psychiatry
	2084P0015X	Physicians/Psychiatry & Neurology, Psychosomatic Medicine
	2084S0010X	Physicians/Psychiatry & Neurology, Sports Medicine
	2084V0102X	Physicians/Psychiatry & Neurology, Vascular Neurology
	2084B0040X	Physicians/Psychiatry & Neurology, Behavioral Neurology & Neuropsychiatry***
	2084S0012X	Physicians/Psychiatry & Neurology, Sleep Medicine***
Radiology/Nuclear Medicine	2085R0001X	Physicians/Radiology, Radiation Oncology
	2085R0202X	Physicians/Radiology, Diagnostic Radiology
	1223X0008X	Oral and Maxillofacial Radiology***
	2085B0100X	Physician/Radiology/Body Imaging***
	2085D0003X	Physician/Radiology/Diagnostic Neuroimaging***
	2085N0700X	Physician/Radiology/Neuroradiology***
	2085N0904X	Physician/Radiology/Nuclear Radiology***
	2085P0229X	Physician/Radiology/Pediatric Radiology***
	2085R0203X	Physician/Radiology/Therapeutic Radiology - Radiation Therapist***
	2085R0204X	Physician/Radiology/Vascular & Interventional Radiology***
	2085R0205X	Physician/Radiology/Radiological Physics***
	2085U0001X	Physician/Radiology/Diagnostic Ultrasound***
	207U00000X	Physicians/Nuclear Medicine***
	207UN0901X	Physicians/Nuclear Medicine, Nuclear Cardiology***
	207UN0902X	Physicians/Nuclear Medicine, Nuclear Imaging & Therapy***
	207UN0903X	Physicians/Nuclear Medicine, In Vivo & In Vitro Nuclear Medicine***
Emergency Medicine	207P00000X	Physicians/Emergency Medicine
	207PE0004X	Physicians/Emergency Medicine, Emergency Medical Services
	207PH0002X	Physicians/Emergency Medicine, Hospice and Palliative Medicine
	207PP0204X	Physicians/Emergency Medicine, Pediatric Emergency Medicine
	207PS0010X	Physicians/Emergency Medicine, Sports Medicine
	207PE0005X	Physicians/Emergency Medicine, Undersea and Hyperbaric Medicine
	207PT0002X	Physicians/Emergency Medicine, Medical Toxicology***
Other	261QM1300X	Ambulatory Health Care Facilities/Clinic/Center, Multi-Specialty
	207RA0001X	Physicians/Advanced Heart Failure and Transplant Cardiology
	207QH0002X	Physicians/Family Medicine, Hospice and Palliative Medicine***
	204C00000X	Physicians/Neuromusculoskeletal Medicine, Sports Medicine***
	207QA0401X	Physicians/Family Medicine, Addiction Medicine***
	207QB0002X	Physicians/Family Medicine, Bariatric Medicine***
	207QS0010X	Physicians/Family Medicine, Sports Medicine***
	207QS1201X	Physicians/Family Medicine, Sleep Medicine***

Table 5.C.6. ***(continued)***

Medicare practitioner-type description	Practitioner taxonomy code	Practitioner taxonomy description
	2080H0002X	Physicians/Pediatrics, Hospice and Palliative Medicine***
	2080N0001X	Physicians/Pediatrics, Neonatal-Perinatal Medicine***
	2080P0201X	Physicians/Pediatrics, Pediatric Allergy & Immunology***
	2080P0202X	Physicians/Pediatrics, Pediatric Cardiology***
	2080P0203X	Physicians/Pediatrics, Pediatric Critical Care Medicine***
	2080P0204X	Physicians/Pediatrics, Pediatric Emergency Medicine***
	2080P0205X	Physicians/Pediatrics, Pediatric Endocrinology***
	2080P0206X	Physicians/Pediatrics, Pediatric Gastroenterology***
	2080P0207X	Physicians/Pediatrics, Pediatric Hematology-Oncology***
	2080P0208X	Physicians/Pediatrics, Pediatric Infectious Diseases***
	2080P0210X	Physicians/Pediatrics, Pediatric Nephrology***
	2080P0214X	Physicians/Pediatrics, Pediatric Pulmonology***
	2080P0216X	Physicians/Pediatrics, Pediatric Rheumatology***
	2080S0010X	Physicians/Pediatrics, Sports Medicine***
	2080S0012X	Physicians/Pediatrics, Sleep Medicine***
	2080T0004X	Physicians/Pediatrics, Pediatric Transplant Hepatology***
	2083A0100X	Physicians/Preventive Medicine, Aerospace Medicine***
	2083P0011X	Physicians/Preventive Medicine, Undersea and Hyperbaric Medicine***
	2083P0500X	Physicians/Preventive Medicine, Preventive Medicine/Occupational Environmental Medicine***
	2083S0010X	Physicians/Preventive Medicine, Sports Medicine***
	2083X0100X	Physicians/Preventive Medicine, Occupational Medicine***
	208VP0000X	Physicians/Pain Medicine, Pain Medicine***
	208VP0014X	Physicians/Pain Medicine, Interventional Pain Medicine***

Source: Centers for Medicare & Medicaid Services. “Crosswalk Medicare Provider/Supplier to Healthcare Provider Taxonomy.” Baltimore, MD: CMS. Available at https://data.cms.gov/Medicare-Enrollment/CROSSWALK-MEDICARE-PROVIDER-SUPPLIER-to-HEALTHCARE/j75i-rw8y. Accessed January 4, 2022.

Notes: Descriptions annotated with three asterisks (***) are categories added since our second annual report. These new specialist categories were added to ensure consistency across measures. The specialty designations are now the same across the measures of ambulatory visits and continuity/fragmentation of care measures.

Table 5.C.7. Ambulatory visit HCPCS/CPT codes and revenue center codes

Place of service	HCPCS/CPT codes	Revenue center codes
Office/outpatient, home; Federally Qualified Health Center; Critical Access Hospital; Rural Health Clinic	99201–99205, 99211–99215, 99324–99328, 99334–99337, 99339–99345, 99347–99350, G2212, 99354–99355, 99358–99359, 99415–99416, 99381–99387; 99391-99397, 98966-98968[a], 99441–99443[a], 98969[a], 99444, 98970–98972, 99421–99423, 99453–99454, 99457, 99458, 99461, 99474, 99483–99484, 99487, 99489–99491, 99439, G2058, G2064–G2065, 99492–99498, G2214, 99091, 90785, 90791–90792, 90832, 90834, 90837, 90833, 90836, 90838–90840, 90845–90847, 90849, 90853, 96150–96155, 96156, 96158, 96164, 96167, 96170, 96159, 96165, 96168, 96171, 99420, 96160–96161, 97151-97158, G0076-G0087, G2010, G2011, G2012, G2061–G2063, G2250-G2252, G2076, G2086–G2088, G0402, G0438, G0439, G0502–G0507, G0513–G0514, G9978–G9986, G9987, 99241–99245[b], 99401–99404[b], 99406[b], 99407–99409[b], 99411–99412[b], 99429[b], G0101–G0102[b], G0108–G0109[b], G0296[b], G0396–G0397[b], G0442–G0447[b], G0473[b], Q0091[b]	n.a.
Federally Qualified Health Center only	G0466–G0468, G0469–G0470	n.a.
Critical Access Hospital only	G0463	
Federally Qualified Health Center or Rural Health Clinic only	G0511, G0512, G0071, G2025	0521, 0522, 0527, 0528

Sources: American Medical Association. “CPT, Professional Edition.” 2016–2021; American Medical Association. “HCPCS Level II, Professional Edition.” 2016–2021.

Note: For this annual report, we expanded the list to include new procedure codes in 2021.

[a] These CPT codes existed prior to 2016 and will not be shown in Table 5.C.9 (code changes instituted by the CPT Editorial Panel during the analytic time period). They were added to the list for the third annual report to align with new online and telephonic assessment and E&M codes the CPT Editorial Panel added in 2019.

[b] These CPT codes existed prior to 2016 and will not be shown in Table 5.C.9 (code changes instituted by the CPT Editorial Panel during the analytic time period). They were added to the list for this report to align with the narrow definition of primary care services that others have used to measure primary care spending in both the Medicare and the commercially insured populations (Bailit et al. 2017; Reid et al. 2019; Kempski and Greiner 2020).

HCPCS/CPT = Health Care Common Procedure Coding System/Current Procedural Terminology; n.a. = not applicable.

Table 5.C.8. Detailed description of the HCPCS/CPT codes and revenue center codes used to identify ambulatory visits

HCPCS/CPT codes	HCPCS/CPT code description	Ambulatory visit indicator[a]	Place of service indicator[b]
99201–99205, 99211–99215	Evaluation and Management (E&M): office or outpatient	1	
99324–99337	Evaluation and Management (E&M): domiciliary, rest home, or custodial care	1	
99339–99340	Evaluation and Management (E&M): domiciliary, rest home, or home care plan oversight	1	
99341–99345, 99347–99350	Evaluation and Management (E&M): home services	1	
G2212	Prolonged office or other outpatient E&M service(s)	0	
99354–99355	Prolonged E&M or Psychotherapy Service w/Direct Patient Contact	0	Yes
99358–99359	Prolonged E&M Service w/o Direct Patient Contact	0	Yes
99415–99416	Prolonged E&M Service w/Direct Patient Contact w/physician supervisor	0	Yes
99381–99387, 99391–99397	Preventive Medicine Services	1	
98966–98968 99441–99443	Telephone assessment & management Telephone E&M	1	
98969 99444	Online assessment & management Online E&M	1	
98970–98972	Online digital assessment	1	
99421–99423	Online digital E&M services – physicians or other qualified health professionals	1	
99453–99454	Chronic Care Remote Patient Monitoring Codes	1	
99457	Remote physiologic monitoring treatment management services, initial 20 minutes	1	
99458	Remote physiologic monitoring treatment management services, additional 20 minutes	0	
99461	Initial care per day, for E&M of normal newborn infant seen in other than hospital or birthing center	1	
99474	Home blood pressure monitoring support	1	
99483	Cognitive Assessment	1	
99484	General Behavioral Health Integration Care Management	1	
99487	Complex Chronic Care Management Services, initial 60 minutes	1	
99489	Complex Chronic Care Management Services , additional 30 minutes	0	
99490	Chronic Care Management, initial 20 minutes	1	
99439	Chronic Care Management, each additional 20 minutes	0	
G2058	Chronic Care Management, each additional 20 minutes	0	
99491	Chronic care management services, provided personally by a physician or other qualified health care professional	1	
G2064	Principal care management service at least 30 minutes	1	
G2065	Principal care management service at least 30 minutes – clinical staff time directed by a physician or other qualified health care professional	1	
99492–99493	Psychiatric Collaborative Care Management (CoCM)	1	
99494	Psychiatric Collaborative Care Management (CoCM), each additional 30 minutes	0	
G2214	Psychiatric Collaborative Care Management	1	
99495–99496	Transitional Care Management Services	1	Yes
99497	Advanced directive counseling and discussion	1	
99498	Advanced directive counseling and discussion, each additional 30 minutes	0	Yes

Table 5.C.8. ***(continued)***

HCPCS/CPT codes	HCPCS/CPT code description	Ambulatory visit indicator[a]	Place of service indicator[b]
99091	Remote Physiologic Patient Monitoring	1	
90785	(Psych) Interactive complexity (in addition to primary procedure)	0	Yes
90791–90792	Psychiatric diagnostic evaluation	1	Yes
90832, 90834, 90837	Psychotherapy	1	Yes
90833, 90836, 90838	Psychotherapy in conjunction w/E&M code	0	Yes
90839	Psychotherapy for crisis	1	Yes
90840	Psychotherapy for crisis, each additional 30 minutes	0	Yes
90845–90847	Other psychotherapy	1	Yes
90849	Multiple family	1	Yes
90853	Group psychotherapy	1	Yes
96150–96151	Health and Behavior Assessment/Intervention	1	Yes
96156	Health behavior assessment or re-assessment	1	Yes
96152–96155	Health & behavior intervention, each 15 minutes	1	Yes
96158, 96164, 96167, 96170	Health behavior intervention, initial 30 minutes	1	Yes
96159, 96165, 96168, 96171	Health behavior intervention, each additional 15 minutes	0	Yes
99420	Administration and interpretation of health risk assessments	1	
96160–96161	Administration of health risk assessment	1	
97151-97158	Adaptive Behavior Therapy assessment and treatment codes	1	
G0076-G0087	Care management home visit	1	
G2010	Remote evaluation of recorded video and/or images submitted by an established patient	1	
G2011	Alcohol and/or substance abuse structured assessment and brief intervention	1	
G2012	Virtual check-in by a physician or other qualified health care professional who can report E&M services	1	
G2061–G2063	Qualified nonphysician healthcare professional online assessment and management service, for an established patient	1	
G2250	Remote assessment of recorded video and/ or images submitted by an established patient	1	
G2251	Brief communication technology-based service, e.g., virtual check-in, by a qualified health care professional who cannot report E&M services	1	
G2252	Brief communication technology-based service, e.g., virtual check-in, by a physician or other qualified health care professional who can report E&M services	1	
G2076	Intake activities, including a physician assessment	1	Yes
G2086–G2088	Office-based treatment for opioid use disorder	1	
G0402	Initial exam for Medicare enrollment	1	
G0438–G0439	Counseling, Wellness, and Screening Services	1	
G0502–G0503	Initial or subsequent psychiatric collaborative care management	1	
G0504	Initial or subsequent psychiatric collaborative care management, each additional 30 minutes	0	
G0505	Cognition and functional assessment using standardized instruments with development of recorded care plan for the patient with cognitive impairment	1	
G0506	Comprehensive assessment and care planning for patients needing chronic care	1	
G0507	Care management services for behavioral health conditions	1	
G0513–G0514	Prolonged Preventive Services	0	
G9978–G9986	Remote in-home visit for the E&M of a patient	1	

Table 5.C.8. ***(continued)***

HCPCS/CPT codes	HCPCS/CPT code description	Ambulatory visit indicator[a]	Place of service indicator[b]
G9987	Bundled payments (BPCI advanced) model home visit for patient assessment	1	
99241–99245	Office or other outpatient consultations	1	
99401–99404	Preventive medicine counseling and/or risk reduction intervention	1	
99406	Smoking and tobacco use cessation counseling visit, greater than 3 minutes up to 10 minutes	1	
99407	Smoking and tobacco use cessation counseling visit, intensive, greater than 10 minutes	1	
99408–99409	Alcohol/Substance Abuse Screening	1	
99411–99412	Group preventive medicine counseling and/or risk reduction intervention	1	
99429	Unlisted preventive medicine service	1	
G0101	Cervical or vaginal cancer screening; pelvic and clinical breast examination	1	
G0102	Prostate cancer screening; digital rectal examination (DRE)	1	
G0108	Diabetes outpatient self-management training services, individual, per 30 minutes	1	
G0109	Diabetes outpatient self-management training services, group session (2 or more), per 30 minutes	1	
G0296	Visit to determine lung cancer screening eligibility	1	
G0396	Alcohol and/or substance abuse structured screening and brief intervention services; 15 to 30 min	1	
G0397	Alcohol and/or substance abuse structured screening and brief intervention services; greater than 30 min	1	
G0442	Annual alcohol misuse screening, 15 minutes	1	
G0443	Brief face-to-face behavioral counseling for alcohol misuse, 15 minutes	1	
G0444	Annual depression screening	1	
G0445	High intensity behavioral counseling to prevent sexually transmitted infection	1	
G0446	Annual, face-to-face intensive behavioral therapy for cardiovascular disease, individual, 15 minutes	1	
G0447	Face-to-face behavioral counseling for obesity, 15 minutes	1	
G0473	Face-to-face behavioral counseling for obesity, group (2–10), 30 minutes	1	
Q0091	Screening Papanicolaou smear; obtaining, preparing and conveyance of cervical or vaginal smear to lab	1	
Critical Access Hospitals only			
G0463	Hospital OP clinic visit	1	
Federally Qualified Health Center only			
G0466–G0467	FQHC visit	1	
G0468	FQHC visit with AWV or IPPE	1	
G0469–G0470	FQHC mental health visit - new patient	1	

Table 5.C.8. ***(continued)***

HCPCS/CPT codes	HCPCS/CPT code description	Ambulatory visit indicator[a]	Place of service indicator[b]
Rural Health Clinic/Federally Qualified Health Center only			
G0071	Non-face-to-face communication between RHC/FQHC practitioner and patient in lieu of an office visit	1	
G0511	General Care Management	1	
G0512	Psychiatric collaborative care management	1	
G2025	Distant site telehealth services	1	

Revenue center codes	Revenue center code description	Ambulatory visit indicator[a]	Place of service indicator[b]
Rural Health Clinic/Federally Qualified Health Center only			
0521	Clinic visit by member to RHC/FQHC	1	
0522	Home visit by RHC/FQHC practitioner	1	
0527	RHC/FQHC Visiting Nurse Service(s) to a member's home when in a home health shortage area	1	
0528	Visit by RHC/FQHC practitioner to other non-RHC/FQHC site (e.g., scene of accident)	1	

Sources: American Medical Association. "CPT, Professional Edition." 2016–2021; American Medical Association. "HCPCS Level II, Professional Edition." 2016–2021.

Notes: This table has been updated to include newly effective codes in 2021. It reflects CPT/HCPCS code changes instituted by the CPT Editorial Panel during the analytic time period (see Table 5.C.9 below). The CPT Editorial Panel comprises 17 members, 11 of whom are physicians, responsible for maintaining the CPT code set for the American Medical Association. Procedure codes used in the identification of non-face-to-face ambulatory visits are shaded in **green**.

[a] Procedure codes with an ambulatory visit indicator of one are included in the visit counts. Indicators with a value of zero indicate add-on services and are not counted as a separate visit.

[b] Some procedure codes that are included in our ambulatory visit definition are also provided in non-ambulatory settings. These services have a place of service indicator equal to "yes" and are counted in our visit and expenditure calculations only if the place of service is not an institutional setting. This excludes services with place of service = 21 (Inpatient Hospital), 51 (Inpatient Psychiatric Facility), 55 (Residential Substance Abuse Treatment Facility), 56 (Psychiatric Residential Treatment Center), or 61 (Comprehensive Inpatient Rehabilitation Facility).

AWV = Annual Wellness Visit; BPCI = Bundled Payments for Care Improvement; CoCM = Collaborative Care Model; E&M = Evaluation and Management; FQHC = Federally Qualified Health Center; HCPCS/CPT = Health Care Common Procedure Coding System/Current Procedural Terminology; IPPE = Initial Preventive Physical Examination; OP = Outpatient; RHC = Rural Health Clinic.

Table 5.C.9. Ambulatory HCPCS/CPT code changes instituted by the CPT Editorial Panel[a] during the analytic time period

HCPCS/CPT codes	HCPCS/CPT code description	Year added	Year replaced
99201	Office or other outpatient visit	Prior to 2016	Deleted in 2021 and not replaced
99444	Online E&M	Prior to 2016	Deleted in 2020 and replaced with 99421–99423
99420	Administration and interpretation of health risk assessments	Prior to 2016	Deleted in 2017 and replaced with 96160–96161
99497	Advance directive counseling and discussion	2016	
99498	Each additional 30 minutes	2016	
96160–96161	Administration of health risk assessment	2017	
98969	Online assessment & management	Prior to 2016	Deleted in 2020 and replaced with 98970–98972
99487	Complex Chronic Care Management Services	2017	
99489	Additional 30 minutes	2017	
99490	Chronic Care Management	2017	
G0502–G0503	Initial or subsequent psychiatric collaborative care management (CoCM)	2017	Deleted in 2018 and replaced with 99492–99494
G0504	Initial or subsequent psychiatric collaborative care management, each additional 30 minutes	2017	Deleted in 2018 and replaced with 99494
G0505	Cognition and functional assessment using standardized instruments with development of recorded care plan for the patient with cognitive impairment	2017	Deleted in 2018 and replaced with 99483
G0506	Comprehensive assessment and care planning for patients needing chronic care	2017	
G0507	Care management services for behavioral health conditions	2017	Deleted in 2018 and replaced with 99484
99091	Remote Physiologic Patient Monitoring	2018	
99483	Cognitive Assessment	2018	
99484	General Behavioral Health Integration Care Management	2018	
99492–99494	Psychiatric Collaborative Care Management	2018	
99453-99454	Chronic Care Remote Patient Monitoring Codes	2019	
99457	Remote physiologic monitoring treatment management services	2019	
99491	Chronic care management services, provided personally by a physician or other qualified health care professional	2019	
97151-97158	Adaptive Behavior Therapy assessment and treatment codes	2019	
G0076- G0087	Care management home visit	2019	
G2010	Remote evaluation of recorded video and/or images submitted by an established patient	2019	
G2011	Alcohol and/or substance abuse structured assessment and brief intervention	2019	
G2012	Virtual check-in by a physician or other qualified health care professional who can report E&M services	2019	
G9978-G9986	Remote in-home visit for the E&M of a patient	2019	
G9987	Bundled payments (BPCI advanced) model home visit for patient assessment	2019	
98970–98972	Online digital assessment	2020	
99421–99423	Online digital E&M services – physicians or other qualified health professionals	2020	
99458	Remote physiologic monitoring treatment management services + 20 minute add-on code	2020	
99474	Home blood pressure monitoring support	2020	

Table 5.C.9. ***(continued)***

HCPCS/CPT codes	HCPCS/CPT code description	Year added	Year replaced
G2058	Chronic Care Management each additional 20 minutes	2020	Deleted in 2021 and replaced with 99439
G2064	Principal care management service at least 30 minutes	2020	
G2065	Principal care management service at least 30 minutes – clinical staff time directed by a physician or other qualified health care professional	2020	
96156	Health behavior assessment or re-assessment	2020	
96158, 96164, 96167, 96170	Health behavior intervention, initial 30 minutes	2020	
96159, 96165, 96168, 96171	Health behavior intervention, each additional 15 minutes	2020	
G2061–G2063	Qualified nonphysician healthcare professional online assessment and management service, for an established patient	2020	
G2076	Intake activities, including a physician assessment	2020	
G2086–G2088	Office-based treatment for opioid use disorder	2020	
99439	Chronic Care Management, each additional 20 minutes	2021	
G2212	Prolonged office or other outpatient E&M service(s)	2021	
G2214	Psychiatric Collaborative Care Management	2021	
G2250	Remote assessment of recorded video and/ or images submitted by an established patient	2021	
G2251	Brief communication technology-based service, e.g., virtual check-in, by a qualified health care professional who cannot report E&M services	2021	
G2252	Brief communication technology-based service, e.g., virtual check-in, by a physician or other qualified health care professional who can report E&M services	2021	
Rural Health Clinic/Federally Qualified Health Center only			
G0511	General Care Management	2018	
G0512	Psychiatric Collaborative Care Management	2018	
G0071	Non-face-to-face communication between RHC/FQHC practitioner and patient in lieu of an office visit	2019	
G2025	Distant site telehealth services	2020	

Sources: American Medical Association. “CPT, Professional Edition.” 2016–2021; American Medical Association. “HCPCS Level II, Professional Edition.” 2016–2021.

[a] The CPT Editorial Panel comprises 17 members, 11 of whom are physicians, responsible for maintaining the CPT code set for the American Medical Association.

BPCI = Bundled Payments for Care Improvement; CoCM = Collaborative Care Model; E&M = Evaluation and Management; HCPCS/CPT = Health Care Common Procedure Coding System/Current Procedural Terminology.

C. Planned care and population health

We constructed a total of 11 claims-based measures under the planned care and population health domain. We constructed six of the measures applying the 2018 specifications obtained from the Healthcare Effectiveness Data and Information Set (HEDIS; available at http://www.ncqa.org/hedis-quality-measurement/hedis-measures/hedis-2018) on Medicare Part A and B claims. The remaining five measures used Part D prescription drug claims data. Two of these were approximations of MIPS clinical quality measures included in the QPP program and were based on measure descriptions from the QPP program; the other three used specifications and value sets from the Pharmacy Quality Alliance (PQA).

C.1. Measures constructed using Medicare Part A and B claims

Five of the six HEDIS measures constructed using Medicare Part A and B claims were for Medicare FFS beneficiaries ages 18 to 75 with diabetes, and one was for breast cancer screening among women ages 52 through 74. In line with the HEDIS specifications, we restricted the five diabetes measures to beneficiaries with continuous Medicare FFS Part A and B enrollment during the 12-month performance period (that is, the year for which the measure is being defined). The breast cancer screening measure required continuous Medicare FFS Part A and Part B enrollment during the 27-month measurement period. Given that we do not have access to more recent versions of the HEDIS specifications, each year we conduct our own review of recent procedure code and diagnosis code changes and update the HEDIS value data sets (VDS) as needed. Our review of new 2021 codes for the VDS for this report identified the following additions to the diabetes measures:

- Added online and telephone assessments at outpatient locations for identifying beneficiaries with diabetes.[30]
- Added two ICD-10 procedure codes to the kidney transplant (nephrectomy) value set.[31]
- Added two new diabetic retinal screening procedure codes.[32]

In Table 5.C.10, we summarize the measure specifications and note where our approach deviates from the approach in the HEDIS specifications. For example, we did not use prescription drug data in constructing these six measures.

[30] The new procedure codes for online assessments are 98969-98972, 99421-99423, 99444, 99457, G0071, G2010, G2012, and G2061-G2063, and the new codes for telephone visits are 98966-98968 and 99441-99443.

[31] The two new procedure codes for kidney transplant are 0TB0 and 0TB1.

[32] The two new retinal screening procedure codes are 92201 and 92202.

Table 5.C.10. Measures based on 2018 HEDIS specifications used for the planned care and population health domain

Measure	Measure numerator	Measure denominator
HbA1c testing	Beneficiaries had an HbA1c test performed during the measurement year.	• Beneficiaries had to be continuously enrolled in FFS Medicare during the measurement year. • Beneficiaries are excluded if they used hospice services during the measurement year. • Beneficiaries ages 18–75 with diabetes (Type 1 or Type 2), defined as having one of the following during the measurement year or the prior year: – Two face-to-face encounters in an outpatient setting or non-acute inpatient setting on different dates of service, with a diagnosis of diabetes. – One face-to-face encounter in an acute inpatient setting, with a diagnosis of diabetes. • Beneficiaries with gestational or steroid-induced diabetes during the measurement year or the prior year were excluded. Notes: We modified the HEDIS “continuously enrolled” criteria by: • Requiring enrollment each month, rather than allowing a 45-day gap in enrollment.(HEDIS considers a beneficiary to have continuous enrollment if the beneficiary had no more than one gap in enrollment of up to 45 days during the measurement year.) • Expanding the criteria for enrollment to match our eligibility criteria for the CPC+ evaluation—a beneficiary is Medicare FFS eligible in a month if the beneficiary is eligible for Part A and Part B with Medicare being the primary payer, not enrolled in an HMO in the month, and alive during any part of the month. We modified the HEDIS denominator by: • Using a broad range of E codes for identification of diabetes diagnoses (E10-E13). • Removing 99420 from the Outpatient VDS (new codes 96160 and 96161 are not included). • Not including code 99483 from the Outpatient VDS.
Eye exam (retinal) performed	Beneficiaries had an eye exam during the measurement year, defined as having one of the following: • A retinal or dilated eye exam by an eye care professional (optometrist or ophthalmologist) in the measurement year. • A negative retinal or dilated eye exam (negative for retinopathy) by an eye care professional in the year prior to the measurement year. Notes: We modified the HEDIS measure by: • Not including eye enucleation in the numerator. • Adding ICD-9 codes for diabetes without complications for prior year identification of retinal exams, because analogous ICD-10 codes were added to the HEDIS measure in 2017.	Same as above

Table 5.C.10. ***(continued)***

Measure	Measure numerator	Measure denominator
Medical attention for nephropathy	Beneficiaries had a nephropathy screening or monitoring test OR evidence of nephropathy during the measurement year, defined as having one of the following during the measurement year: • A nephropathy screening or monitoring test • Evidence of treatment for nephropathy or ACE/ARB therapy • Evidence of Stage 4 chronic kidney disease • Evidence of end-stage renal disease • Evidence of kidney transplant • A visit with a nephrologist	Same as above
Composite diabetes care measure for receiving all three tests	Beneficiaries received all three tests during the measurement year—an HbA1c test, an eye exam, and medical attention for nephropathy.	Same as above
Composite diabetes care measure for not receiving any of the three tests	Beneficiaries did not receive any of the three tests during the measurement year—an HbA1c test, an eye exam, and medical attention for nephropathy.	Same as above
Breast cancer screening	Beneficiaries with one or more mammograms any time on or between October 1 two years prior to the start of the measurement year and December 31 of the measurement year.	• Beneficiaries had to be continuously enrolled during the measurement year and for the 15 months prior to the measurement year. • Beneficiaries are excluded if they used hospice services during the measurement year. • Women ages 52–74 as of December 31 of the measurement year. • Beneficiaries who had a bilateral mastectomy or a right and a left unilateral mastectomy were excluded. We used claims back to 2013 to identify these exclusions. Note: • This measure incorporated the same deviations from HEDIS for the continuously enrolled criteria.

Source: National Committee for Quality Assurance (NCQA). "HEDIS Volume 2: Technical Specifications." 2018.

HbA1c = Hemoglobin A1c test; HEDIS = Healthcare Effectiveness Data and Information Set; ICD-9 = International Classification of Diseases Version 9; ICD-10 = International Classification of Diseases Version 10; VDS = HEDIS value data set.

C.2. Measures constructed using Medicare Part D claims

We created two measures that were approximations of MIPS clinical quality measures included in the QPP program: (1) Percentage of beneficiaries with cardiovascular disease who were prescribed statin therapy ("statin therapy") and (2) Percentage of beneficiaries with both coronary artery disease (CAD) and diabetes who were prescribed angiotensin-converting enzyme (ACE) inhibitors or angiotensin receptor blockers (ARB) therapy ("ACE/ARB therapy"). These measures were restricted to beneficiaries who had continuous Medicare FFS Parts A, B, and D enrollment during the measurement year and no hospice utilization that year. Table 5.C.11 provides details on the denominators and numerators for these measures.

Table 5.C.11. Prescription drug-related measures based on 2021 MIPS specifications used for the planned care and population health domain

Measure	Measure numerator	Measure denominator
Percentage of beneficiaries with cardiovascular disease who were prescribed statin therapy	Receipt of a statin medication as identified in the Part D prescription drug event data during the performance year	• Beneficiaries had to be continuously enrolled in FFS Medicare and Medicare Part D during the measurement year. • Beneficiaries 21 years of age or older who were previously diagnosed with or currently have an active diagnosis of atherosclerotic cardiovascular disease or who have an active diagnosis of familial or pure hypercholesterolemia during the measurement year. • Beneficiaries are excluded if they used hospice services, were pregnant or breastfeeding, or had a diagnosis of rhabdomyolysis during the measurement year. Exceptions include active liver or hepatic disease or insufficiency or end-stage renal disease.
Percentage of beneficiaries with both CAD and diabetes who were prescribed ACE inhibitors or ARB therapy	Receipt of an ACE/ARB medication as identified in the Part D prescription drug event data during the performance year	• Beneficiaries had to be continuously enrolled in FFS Medicare and Medicare Part D during the measurement year. • Beneficiaries 18 years old or older with two encounters with diagnoses of CAD and diabetes during the measurement year. • Beneficiaries are excluded if they used hospice services during the measurement year.

Notes: Yearly NDC mappings from NCQA were used to identify ACE/ARB medications. (The downloadable NDC files are available in the HEDIS® technical resources section at https://www.ncqa.org/hedis/measures/.)

We expanded the criteria for Medicare FFS enrollment to match our eligibility criteria for the CPC+ evaluation—beneficiaries are Medicare FFS eligible in a month if they are enrolled in both Part A and Part B with Medicare being the primary payer, not enrolled in an HMO during the month, and alive during any part of the month.

ACE = angiotensin-converting enzyme; ARB = angiotensin receptor blockers; CAD = coronary artery disease; FFS = fee-for-service; HMO = health maintenance organization; NCQA = National Committee on Quality Assurance; NDC = National Drug Codes.

Percentage of beneficiaries with cardiovascular disease who were prescribed statin therapy. The statin therapy measure approximates the MIPS clinical quality measure "statin therapy for the prevention and treatment of cardiovascular disease" (Quality ID #438). Because we cannot measure the concept of prevention or determine low-density lipoprotein cholesterol levels in claims data, the denominator for our approximation is restricted to adults 21 years of age or older who were previously diagnosed with or currently have an active diagnosis of atherosclerotic cardiovascular disease or who have an active diagnosis of familial or pure hypercholesterolemia. (A detailed description of the 2021 MIPS measure can be found at Quality ID #438: Statin Therapy for the Prevention and Treatment of Cardiovascular Disease [https://qpp.cms.gov/docs/QPP_quality_measure_specifications/CQM-Measures/2021_Measure_438_MIPSCQM.pdf].)

Percentage of beneficiaries with both CAD and diabetes who were prescribed ACE inhibitors or ARB therapy. The ACE/ARB therapy measure approximates the MIPS clinical quality measure "coronary artery disease: ACE inhibitor or ARB therapy - diabetes or left ventricular systolic dysfunction (LVEF < 40%)" (Quality ID #118). The denominator for our approximation is restricted to beneficiaries 18 years old or older

with CAD and diabetes because we cannot identify LVEF in claims data. (A detailed description of the 2021 MIPS measure can be found at https://qpp.cms.gov/docs/QPP_quality_measure_specifications/CQM-Measures/2021_Measure_118_MIPSCQM.pdf.)

Finally, we constructed three measures using specifications and value sets from the Pharmacy Quality Alliance (PQA). These measures are the percentage of beneficiaries on diabetes medications, renin-angiotensin system antagonists, or statins, respectively, with proportion of days covered by medication > 80%.

The denominator for each measure is beneficiaries 18 years or older with at least two dispensing events for a qualifying medication during the year, where a dispensing event is defined as a record in the Part D event data indicating the medication was dispensed by a pharmacy. These measures were restricted to beneficiaries with continuous Medicare FFS Part A, B, and D enrollment during the measurement year who had a treatment period of at least 91 days[33] (that is, the number of eligible days, defined below, was greater than 91). Denominator exclusions are receipt of hospice care or diagnosis of end-stage renal disease (ESRD) during the measurement year. For the diabetes measure, receipt of insulin as identified in the Part D prescription drug event data is an additional exclusion.

The first step to construct the numerators is to determine the number of eligible days, which is the number of days from the first dispensing event to the end of the measurement year for those who did not die during the year, or the number of days from the first dispensing event to the date of death for those who died during the year. Next, the number of days' supply is calculated from all the dispensing events identified in the Part D prescription drug event data during the measurement year. We account for overlapping days' supply in the following manner:

- **Overlap of 14 days or fewer.** We consider this to be an early refill and we add the day supply amount to the overall count of days. For example, if fill 1 occurred on June 1, 2020, a 90-day supply would end on August 29, 2020. If a second dispensing event occurred on August 25, 2020, with 90 days' supply, we would add the days' supply from the two separate dispensing events and count these two events as having a total of 180 days.
- **Overlap of more than 14 days and the next dispensing date is the same as the current dispensing date for different medications in the same class.** This suggests a complementary medication regimen, so we use the dispensing event with the maximum number of days to set the days' supply. For example, if a 90-day supply of medication 1 was dispensed on June 1, 2020, and a 30-day supply of medication 2, which is in the same class as medication 1, was dispensed on the same date, the total days' supply is 90.
- **Overlap of more than 14 days and the next dispensing date is after the current dispensing date.** This could indicate a switch in medications, so we add the days count from the second dispensing event to the number of days from the previous fill through the date of the second dispensing event. For example, if a 90-day supply was dispensed on June 1, 2020, and another 90-day supply was dispensed on July 1, 2020, the total days' supply would be 121 (31 days from the first dispensing event and 90 days from the second dispensing event).
- **Complete overlap with the previous dispensing event.** The days' supply from the two events are not added and we use the days' supply from one event in the measure calculation.

[33] PQA added this eligibility requirement in their measurement year 2021 specifications.

To construct the final measure, we divide the number of days' supply by the number of eligible days. If the result is greater than 0.80, then the beneficiary is considered numerator compliant. We repeat this process for each of the three medications to produce three binary indicators of compliance for the outcome variables. (More information about the PQA measures is available at https://www.pqaalliance.org/assets/Measures/PQA_Measures_Overview.pdf.)

D. Continuity of care

We created five outcomes measures to examine continuity of care, and we describe those measures in greater detail below. The first one is based on ambulatory visits with primary care practitioners (defined earlier in Section B.4) at a beneficiary's assigned practice The next four (two different versions of percentage of visits with the usual provider of care [UPC] and Reverse Bice-Boxerman Index [rBBI]) are based on a slightly narrower set of ambulatory visits to both primary care and specialist practitioners (we refer to these as "qualifying visits") and measure the percentage of those visits with the most frequently seen practitioner and the dispersion of those visits across all practitioners. Beneficiaries were required to meet three criteria to be included in the percentage of visits with the UPC and rBBI continuity of care measures: (1) be in the intent-to-treat (ITT) sample at the beginning of the year; (2) be enrolled in Medicare FFS for the full year; and (3) receive qualifying ambulatory visits in the measurement year.

Percentage of primary care ambulatory visits provided at a beneficiary's assigned practice. For the beneficiaries we identified as having ambulatory visits (Table 5.C.7) with a primary care practitioner (Table 5.C.5), we further examined the percentage of primary care ambulatory visits that were provided by practitioners affiliated with the beneficiary's assigned practice.

In this report, we created two versions of the additional continuity of care measures. The first counts each practitioner individually. Since fragmentation calculated at the practitioner (NPI) level may overstate true fragmentation when there is team-based care, we created a second version of the UPC and rBBI measures that combined practitioners in a beneficiary's assigned primary care practice. All practitioners (NPIs) affiliated with a beneficiary's assigned practice were counted as one practitioner instead of being counted as individual practitioners.

Percentage of visits with the usual provider of care where each practitioner is counted separately. The percentage of visits with the UPC measures the proportion of qualifying ambulatory visits with the most frequently seen ambulatory practitioner (Breslau and Reeb 1975; Pollack et al. 2016). Note that the most frequently seen practitioner could have any specialty (e.g., primary care or specialist). UPC was created for beneficiaries with one or more qualifying ambulatory visits. We used a modified version of the National Committee for Quality Assurance's definition of ambulatory visits to identify beneficiaries with office or other outpatient visits (such as to rural health clinics and critical access hospitals) for E&M; ophthalmological services for medical examination and evaluation; or new enrollee and annual wellness visits (Kern et al. 2017; NCQA 2015). A description of these visit codes can be found in Table 5.C.12. The formula for the measure is:

$$\max\left(\frac{n_i}{N}\right) \text{over all } i \text{ practitioners}$$

Where n_i is the number of ambulatory visits to practitioner i (NPI) during the measurement period, and N is the total number of all ambulatory visits the beneficiary had during the measurement period.

Reversed Bice-Boxerman Index where each practitioner is counted separately. The Bice-Boxerman Continuity of Care Index (COCI) identifies the number of practitioners providing ambulatory services to a beneficiary and the percentage of care provided by each practitioner. The index is created for each beneficiary and is calculated by taking the number of visits to each individual practitioner divided by the total number of visits the beneficiary had overall. A description of the qualifying ambulatory visits is found in Table 5.C.12. This index weights both the frequency of ambulatory visits to each practitioner and the dispersion of visits between practitioners. Index values range from just greater than 0 (visits made to many practitioners) to 1 (all visits made to the same practitioner).

BBI is defined as

$$\left(\sum n_i^2 - N\right) / \left[N\left(N-1\right)\right],$$

where n_i is the number of visits that the beneficiary had with the $i\text{th}$ practitioner, and N is the total number of all ambulatory visits the beneficiary had during the measurement period.

We required beneficiaries to have at least four ambulatory visits to qualify for inclusion in the rBBI, because measures of continuity may not be reliable if they are based on three or fewer visits (Nyweide et al. 2013). To measure fragmentation, we reversed raw BBI scores, calculating 1 minus BBI, for beneficiaries who had at least four ambulatory visits. On this rBBI index, higher scores reflect more fragmentation (many providers with a relatively low proportion of ambulatory visits by each provider). Thus, beneficiaries with an rBBI of 0 have no fragmentation of care (all their qualifying visits were to the same provider).

Measuring both the UPC and rBBI is useful, because the UPC facilitates interpretation. Measuring the percentage of visits with the UPC alongside the rBBI can make the findings more transparent, as the difference between two UPC scores (e.g., 30 percent of visits vs. 50 percent of visits with the most frequently seen provider) is easier to interpret than the clinical difference between two rBBI scores (e.g., 0.9 vs. 0.7).

Percentage of visits with the usual provider of care and Reversed Bice-Boxerman Index where all practitioners at the beneficiary's assigned practice are counted as one practitioner. These two outcomes are defined the same as those above, except that all NPIs associated with the beneficiary's assigned practice are counted as a single practitioner.

Table 5.C.12. Procedure codes used for the selection of qualifying ambulatory visits for the UPC and rBBI measures

HCPCS/CPT codes	Description
99201-99205; 99211-99215	Office or other outpatient visit for E&M
92002, 92004, 92012, 92014	Ophthalmological services: medical examination and evaluation
G0402, G0438, G0439	New enrollee and annual wellness visits

E&M = Evaluation and Management.

E. Comprehensiveness of care

We developed three NPI-level measures intended to gauge the comprehensiveness of care provided by primary care physicians. These measures are slight modifications of those originally developed by O'Malley et al. (2019) and Rich et al. (2021). Comprehensiveness is the extent to which a primary care physician meets the large majority of their patient's physical and common mental health care needs. These measures are created for primary care physicians only. Thus, the measures exclude approximately one-third of CPC+ and comparison group providers because they are nurse practitioners, physician assistants, or physician specialists.[34] We identify a primary care physician based on the physician's NPI in the Medicare Data on Provider Practice and Specialty (MD-PPAS) file being assigned to a taxonomy code in one of the following specialties: 01 (general practice), 08 (family practice), 11 (internal medicine), 37 (pediatric medicine), or 38 (geriatric medicine). To ensure findings on changes associated with CPC+ are not driven by changes in the set of physicians included in the analysis, we limited the set of physicians to those eligible for the analysis in 2017.[35] We describe the development of these measures here.

Involvement in patient conditions (IPC). For each physician, this measure calculates the percentage of beneficiaries seen in a given year (2016-2021) for whom the physician had the greatest involvement in the patient's conditions. To be included in the analysis, a beneficiary must be eligible for Part A and Part B with Medicare being the primary payer, not enrolled in an HMO, and alive during any part of the analysis period. To calculate this measure, we first identify all beneficiaries seen by a CPC+ or comparison group primary care physician in a given year. We identify all the diagnoses for which the beneficiaries were seen by any physician (both primary care and specialists) for an office-based or telehealth E&M service, truncated to the first four digits for ICD-10 codes, and we count the total number of these unique diagnosis codes in office-based E&M services (99201 to 99205, 99211 to 99215) and telehealth E&M services using both CPT and HCPCS codes (99421–99423, 99441–99443, 99444; G2010, G2012, G2061–G2063). Once we identify the set of claims with the CPT and HCPCS codes for each physician and beneficiary combination, we count the total number of the beneficiary's unique diagnoses on these claims for which the physician billed in the year. We look across the physicians who treated the beneficiary, identify the physician who billed for the plurality of the beneficiary's diagnosis codes, and assign that physician as the most comprehensive for that beneficiary. If multiple physicians billed for the

[34] We estimated the comprehensiveness of primary care physicians rather than nurse practitioners (NPs) and physician assistants (PAs), because of the low prevalence of NPs and PAs serving as a patient's usual practitioner in our sample, and the difficulty of discerning all services independently provided by NPs/PAs because they commonly bill "incident to" services under a physician's NPI.

[35] For the CPC+ years, we only include NPIs eligible for the analysis in 2017, holding the set of NPIs and their practice affiliation fixed at their 2017 practice as long as they are observed in SK&A/OneKey data (we do not allow any new NPIs to enter the sample in 2018, 2019, etc.). We do drop 2017 NPIs if they are no longer in SK&A/OneKey data in a later year. All eligible NPIs are included for 2016 even if they were not eligible in 2017.

same share of a beneficiary's diagnoses, then we designate all those physicians as the most comprehensive for that beneficiary. Finally, for each physician, we calculate the share of the beneficiaries treated by the physician for whom the physician was the most comprehensive physician.

New problem management (NPM). This measure assesses the extent to which a physician manages a patient's new symptom or problem instead of referring them to (or the patient seeking) a specialist. The measure focuses on management of the 20 most common reasons for visits to primary care in the Medicare population aged 65 and over.[36]

We calculate this measure annually. For each year, for each beneficiary receiving office-based or telehealth E&M services from a CPC+ or comparison group primary care physician based on the performing physician's NPI, we select the first claim for these services with each condition in Table 5.C.13 based on the diagnosis codes associated with the condition. We call this the index claim for the beneficiary and condition in the analysis year. We exclude index claims for beneficiaries who are not eligible for the analysis for at least 20 months in the 24 months prior to the index claim thru date and for at least 10 months of the 12 months following the index claim through date. To be eligible for the analysis in a particular month, a beneficiary must be eligible for Part A and Part B with Medicare being the primary payer, not enrolled in an HMO, and alive during any part of the month. Because we want to analyze only "new" problems, we also exclude index claims for which the beneficiary had the same diagnosis on any E&M service[37] performed by any provider in the 24 months prior to the index claim "thru date." We define office-based and telehealth E&M services to include all codes listed in Table 5.C.14. After these exclusions, we end up with an output file including index claims for all beneficiaries who saw a CPC+ or comparison practice physician for a "new" condition in the year. Next, for each index claim, we identify all office-based and telehealth E&M services with the same beneficiary and condition in the 12 months following the thru date of the index claim and use these claims to calculate the index physician's share of claims for the "new" condition. Then, separately for each of the 20 conditions, we calculate the average share of services performed by the index claim physician. The calculation of this average includes all physicians with an index claim for the condition. Finally, for each physician, we calculate a new problem management score. This measure assesses the extent to which a physician manages a patient's new symptom or problem instead of referring them to (or the patient seeking) a specialist. We calculate the average share of services the physician provided in the following 12 months for all their "new" condition index claims.[38] To account for differences across physicians in the mix of conditions, we also calculate the predicted value, which is the average of the physician averages with the same mix of conditions. We calculate the new problem management as the ratio of the physician's own average and the predicted average.

[36] The 20 most common reasons for visits to primary care in the Medicare FFS population aged 65 and older are migraine, headache, urinary tract infection, gastrointestinal symptoms, skin disorders, back problems, hypertension, hyperlipidemia, diabetes, depression, anxiety, arthritis and localized joint syndromes, obesity, asthma, ill-defined conditions, upper respiratory conditions, ischemic heart disease, congestive heart failure, chronic obstructive pulmonary disease, and thyroid disorders.

[37] The third annual report used Berenson-Eggers Type of Service (BETOS) codes to define E&M services. We shifted to using CPT/HCPCS codes to identify E&M services for the fifth annual report. Diagnosis codes are used to identify the set of services associated with each condition. The set of diagnosis codes used has been updated to address coding changes since the third annual report.

[38] This measure is only reported for 2016 through 2020 to accommodate the 12 month look-forward period.

Table 5.C.13. Diagnosis codes for new problem management measure

Condition	ICD-9 codes[a]	ICD-10 codes[b]
Migraine	346	G43
Headache	7840	G441, R51
Urinary tract infection	5990	N390
Gastrointestinal symptoms—includes GERD, acute gastritis without hemorrhage, infectious colitis, enteritis, and gastroenteritis, salmonella gastroenteritis	0030, 0090, 0091, 53011, 53012, 5589, 578	A020, A09, K209, K210, K523, K5283, K5289, K529, K920-K922
Skin disorders	680-709	B781, E08628, E09628, E832, I7023-I7025, I7033-I7035, I7043-I7045, I7053-I7055, I7063-I7065, I7073-I7075, K122, L00-L05, L080, L088, L10-L14, L20-L30, L40-L43, L440-L443, L448, L449, L45, L49-L60, L62-L68, L70-L75, L80-L88, L89000-L89004, L89009-L89014, L89019-L89024, L89029, L89100-L89104, L89109-L89114, L89119-L89124, L89129-L89134, L89139-L89144, L89149-L89154, L89159, L89200-L89204, L89209-L89214, L89219-L89224, L89229-L89304, L89309-L89314, L89319-L89324, L89329, L8940-L8945, L89500-L89504, L89509-L89514, L89519-L89524, L89529, L89600-L89604, L89609-L89614, L89619-L89629, L89810-L89814, L89819 L89890-L89894, L89899, L8990-L8995, L90-L93, L940-L945, L948 L949, L95, L97-L99
Back problems (new onset low back pain)	724	M432, M438X9, M4800, M4804-M4808, M532X7, M532X8, M533, M5380, M5384-M5388, M539 M5403-M5409, M5414-M5417, M543-M546, M5489, M549, M62830, M9922-M9929, M9932-M9939, M9942-M9949, M9952-M9959, M9962-M9969, M9972-M9979
Hypertension	401	I10, I160, I161, I169
Hyperlipidemia, lipid disorders	272	E7130, E7521, E7522, E7524, E753, E755, E756, E770, E771, E7841, E7849, E778-E786, E7870, E7879, E788, E789, E881, E882, E8889
Diabetes	249-250	E08-E11, E13
Depression	296.2, 311, 309	F320-F325, F329, F431, F432, F438, F439, F930, F948
Anxiety	300	F341, F40, F41, F42, F422, F423, F428, F429, F44, F450-F452, F458, F459, F481, F488, F489, F6811, F6813, F688, F99, R452, R455, R456
Arthritis and localized joint syndromes	710-716	A1801, A1802, A5216, E08610, E08618, E09610, E09618, E106, E116, E136, M00-M02, M042, M048, M049, M05-M07, M080, M082, M083, M084, M088, M089, M11, M120, M121, M125, M128, M129, M13-M19, M32-M34, M350, M351, M352, M355, M358, M359, M36
Obesity	278	E65, E6601, E6609, E661, E662, E663, E668, E669, E670, E671, E672, E673, E678, E68
Asthma	493	J440, J441, J449, J4520, J4521, J4522, J4530, J4531, J4532, J4540, J4541, J4542, J4550, J4551, J4552, J45901, J45902, J45909, J45990, J45991, J45998
Symptoms, signs, and ill-defined conditions	780–799, except 7840 (7840 is used for headache)	B349, E035, E0781, E0852, E0952, E1052, E1152, E1352, E790, G4700, G4710, G4730, G479, G933, I7036, I7046, I7056, I7066, I7076, I7301, I96, K522, K5229, K5289, N23, N393, N394, O28, P09, R000, R002, R008, R009, R01, R03-R05, R0600-R0602, R0609, R061-R069, R07, R090, R092, R093, R0982, R0989, R10, R110, R1110-R1112, R1114, R1115, R112, R12, R13-R23, R25, R260, R261, R2681, R2689, R269, R27, R290-R293, R295, R296, R298, R299, R30, R32-R35, R360, R369, R39, R400, R401, R4020, R40211, R40212, R40221, R40222, R40231, R40232, R40234, R403, R404, R410-R414, R4181, R4182, R4184, R4189, R419, R42, R43, R440, R442-R449, R450, R453, R454, R4583, R4584, R4586-R4589, R46, R47, R480-R482, R488, R489, R49, R50, R52-R57, R59-R64, R6521, R680, R681, R683, R688, R69-R71, R73-R79, R800, R801, R803, R808, R809, R81-R94, R97, R99, R828, R8281, R8289, R8299, R938

Table 5.C.13. ***(continued)***

Condition	ICD-9 codes[a]	ICD-10 codes[b]
Upper respiratory conditions (not including asthma)	460–477	J00, J01, J028, J029, J038, J039, J04-J06, J20, J21, J30-J33, J342, J35-J37
Ischemic heart disease	413, 414	I201, I208, I209, I251, I253, I2541, I2542, I255, I256, I2570-I2573, I2575, I2576, I2579, I2581-I2584, I2589, I259
CHF	428	I50
Obstructive airway diseases or COPD, asthma	491	J41, J42, J44
Thyroid disorder	246	E034, E041, E070, E071, E0789, E079, E35

Source: American Medical Association. "ICD-10-CM: The Complete Official Codebook." 2015–2021.

[a] We include all ICD-9 codes that *start with* these codes. ICD-9 codes were used for Medicare billing prior to October 1, 2015. They were needed in this analysis to identify whether the beneficiary had the same diagnosis on any E&M service in the 24 months prior to the index claim.

[b] We include all ICD-10 codes that *start with* these codes. ICD-10 codes were used for Medicare billing starting October 1, 2015.

CHF = congestive heart failure; COPD =chronic obstructive pulmonary disease; E&M = evaluation and management; GERD = gastroesophageal reflux disease.

Table 5.C.14. Procedure codes used to identify E&M services for new problem management measure

Category	CPT/HCPCS codes
Office- and non-office-based E&M codes	G0068-G0070, G0076-G0087, G0101, G0245-G0248, G0250, G0378-G0384, G0402, G0420, G0421, G0463, G0466-G0470, G0473, G0490, G2001-G2009, G2011, G2013-G2015, G2082, G2083, G9978-G9986, 0500F, 0502F, 0503F, 1000F, 2000F, 94002-94005, 94660, 94662, 95115, 95117, 99026, 99027, 99058, 99175, 99201-99205, 99211-99215, 99217-99226, 99231-99236, 99238, 99239, 99281-99285, 99288, 99291, 99292, 99304-99310, 99315, 99316, 99318, 99324-99328, 99334-99337, 99341-99345, 99347-99350, 99354-99357, 99366, 99367, 99381-99387, 99391-99397, 99401-99404, 99411, 99412, 99420, 99429, 99466-99469, 99471, 99472, 99475, 99476, 99480, 99485, 99600-99602, 99415, 99416, 99484, 99490, 99491, 99492-99494, 99497, 99498, G0438, and G0439
Telehealth codes	99421-99423, 99441-99444, G0406-G0408, G0425-G0427, G0508-G0509, G2010, G2012, G2061-G2063

Sources: American Medical Association. "CPT, Professional Edition." 2016–2021; American Medical Association. "HCPCS Level II, Professional Edition." 2016–2021.

E&M = evaluation and management.

Range of services (ROS). This measure assesses the range of services a primary care physician provided to their Medicare patients by counting the number of the following types of services the physician provided: immunization administration, behavioral or mental health counseling, cryotherapy/skin excisions, joint injections, and treatment of minor lacerations.

We calculate this measure for each calendar year from 2016 to 2019.[39] For each CPC+ or comparison group primary care physician, we create five indicator variables, one for each type of service represented in the measure. The indicators for a physician are set to 1 if we identify one or more Medicare Part B claims with a date of service during the measurement year with the physician's NPI listed as the performing physician and at least one of the CPT codes listed in Table 5.C.15 for the respective type of service. The indicators are summed to create a final ROS score from 0 (physician did not provide any of the types of service) to 5 (physician provided all of the types of service) for each measurement year.

Table 5.C.15. Procedure codes used to identify select service types for ROS measure

Type of service	CPT/HCPCS codes
Immunization administration	90471, 90472, G0008, G0009
Behavioral or mental health counseling	90791, 90792, 90832-90834, 90836-90838, 90853, 99484, 99492-99494, G0502, G0503, G0504[a]
Treatment of minor lacerations	12001, 12002, 12004, 12005, 12011, 12013, 12014, 12020, 12021, 12031, 12032
Cryotherapy/skin excision	10060, 10061, 10160, 11100, 11101, 11102-11107[b], 11300-11303, 11305-11307, 11310-11312, 11400, 11401-11404, 11420-11422, 11440-11442, 17110, 17250
Joint injection	20550, 20551, 20600, 20605, 20610

Sources: American Medical Association. "CPT, Professional Edition." 2016–2021; American Medical Association. "HCPCS Level II, Professional Edition." 2016–2021.

ROS = range of services.

[a] Behavioral or mental health counseling codes G0502, G0503, and G0504 were added for 2017, but then replaced by 99491-99494 in 2018.

[b] Cryotherapy/skin excision codes 11100 and 11101 were deleted in 2019 and replaced by six new codes (11102–11107) that are based on the thickness of the sample and the technique used for skin excision.

F. Patient and caregiver experience

We created three measures of hospice service use to measure patient and caregiver experience: percentage of beneficiaries using hospice service; days of hospice use for beneficiaries receiving hospice services; and days of hospice use for all beneficiaries.

Any use of hospice services. This measure is the percentage of beneficiaries who received any hospice services in the year. Beneficiaries are identified as having hospice services if they have a hospice claim in the year.

Number of days of hospice use among beneficiaries who received any hospice service during the year. This measure is the number of days a beneficiary spent in hospice care in a given year including days that were reported on denied claims when these claims did not overlap with dates of service on approved claims. We include denied claims to comprehensively account for the services beneficiaries received. To identify the number days of hospice care, we sorted hospice claims by beneficiary

[39] Because these services involve face-to-face visits, this version of the ROS measure is currently being tested for validity during the pandemic. Thus, we are not reporting it beyond 2019. The new measure being tested adds some new services now being provided by some primary care physicians. A separate memo will be submitted once analyses are complete documenting whether this updated measure provides more meaningful information on primary care physician comprehensiveness.

identification number, from date, and through date. Next, we combined claims with overlapping dates of service into a single span of service. Then, we calculated the days in each span by calculating the difference between the through date and the from date on the span and adding one. Finally, for each beneficiary and month, we summed the days in the spans with through dates in the month.

Number of days of hospice use among all beneficiaries. This measure is the number of hospice days in the measurement year, regardless of whether a beneficiary received any hospice services.

G. Other quality of care

We examined 7 additional quality of care outcomes that are based on use of Medicare services. There are three discharge-level measures: unplanned 30-day readmissions, unplanned acute care following an acute hospital discharge, and unplanned acute care following a discharge from an ED. One measure assessing the use of high-risk medications in the elderly. In addition, we have two measures looking at the use of opioids (long-term opioid use and potential opioid overuse) and a measure of low-value service utilization. We describe these measures in more detail below.

Unplanned readmissions within 30 days of a hospital index discharge. For calculating the 30-day readmission rate, we used a slightly different time period definition than for the other measures. We looked at all eligible inpatient discharges during the last month of the previous year and the first 11 months of the current year,[40] and calculated the proportion of these index discharges that were followed by an unplanned hospitalization within 30 days of the discharge. An unplanned readmission is defined as any hospitalization that does not continue care (examples of planned admissions include recurring admissions for chemotherapy and planned admission for transplant surgery).

For an index discharge to qualify for inclusion in the readmission measure, the beneficiary must (1) be enrolled in Medicare FFS Part A and not in a health maintenance organization (HMO) at the time of the index admission, (2) be enrolled in Medicare FFS Part A during the month following discharge, (3) be alive at discharge, and (4) not be discharged against medical advice. In addition, certain inpatient stays were excluded from the universe of index discharges, including discharges with lengths of stay longer than one year; stays at cancer hospitals exempt from the Prospective Payment System; and stays for psychiatric conditions, rehabilitation, or cancer. Our definition of this measure is based on the Yale readmission measure developed by the Yale New Haven Health Services Corporation/Center for

[40] We examine all index discharges during the last month of the previous year and the first 11 months of the current year to ensure that the relevant outcome "readmission within 30 days" is observed within the analysis period with adequate claims runout. One minor disadvantage is that, for the first intervention year, some readmissions are measured in the last month of the baseline (December 2016), before the CPC+ intervention began, which would dilute any observed effect on readmissions in Year 1. However, this factor affects only 1 out of 13 months (12 months of index discharges plus one additional month to observe 30 day readmissions post index discharge) of observed readmissions in Year 1, and should not discernibly change the Year 1 effect, especially because we do not expect the intervention to have sizable effects in Year 1. We considered the alternative of including index discharges over all 12 months of a calendar year. However with this approach, we would not be able to observe all possible 30-day readmissions without expanding the analysis period into the first month of the following year, which for the fifth year of CPC+ would include a month after the intervention ended. Also, it would lead to limited claims runout of only two months for that last month of readmissions in each measurement period.

Outcomes Research & Evaluation (YNHHSC/CORE 2021) that is used in the Hospital Readmission Reduction Program under Section 3025 of the Affordable Care Act.[41]

After we identify the index discharge and qualifying readmissions, we apply these beneficiary eligibility criteria to the readmission: (1) enrolled in Medicare Part B with Medicare as the primary payer in the month of the admission and the month following the admission and (2) enrolled in Medicare Parts A and B, not in an HMO, with Medicare as the primary payer in the month of the discharge. If beneficiaries did not meet these criteria, we did not include them in our readmission measure.

Although we analyze our main readmission outcome at the discharge level, we also conduct a sensitivity test examining the measure of unplanned readmission at the beneficiary level (for motivation and details, see Appendix 5.C). Unlike the discharge level outcome, all beneficiaries in the ITT sample are included in the beneficiary-level analysis. This binary measure takes the value 1 if the beneficiary had a qualifying readmission in the observation period (after applying the eligibility criteria, as explained above), and 0 otherwise.

Unplanned acute care. We developed two binary measures of unplanned acute care based on:

1. Percentage of index acute care hospital discharges that were followed by an unplanned acute care hospitalization or ED visit (including observation stays) within 30 days.
2. Percentage of index ED (including observation stay) discharges that were followed by an unplanned acute care hospitalization or ED visit (including observation stays) within 30 days.

The purpose of these measures is to capture additional unplanned acute care use beyond the 30-day unplanned readmission measure.

To develop the first measure, we start with the set of index hospitalizations used to calculate the 30-day unplanned readmission measure for each measurement year. This is the denominator for the measure. Then, we identify ED discharges (including observation stays) that started within 30 days of the discharge date of the index hospitalization. If the index hospitalization had an unplanned hospital readmission, an ED visit, or an observation stay within 30 days following the index discharge date, we flag the index hospitalization as being followed by unplanned acute care use within 30 days.

To develop the second measure, we first identify all ED visits (including observation stays) with a discharge date in January through November of the measurement year and in December of the prior year. We combine the visits that begin on the same day into one event. We consider these the set of index ED discharges for the measure denominator. Next, we obtain the set of unplanned hospital stays developed for the 30-day unplanned readmission measure for the measurement year and identify those that have an admission date within 30 days of an index ED visit. Then we identify ED visits (including observation stays) that started within 30 days of one of the index ED visits. We flag index ED visits as being followed by unplanned acute care use if they had either an unplanned hospital stay or an ED visit (including observation stays) within 30 days of the index ED visit discharge date.

[41] Additional information about the Yale readmission measure is available at QualityNet, "Measure Methodology Reports: Readmissions Measures," https://qualitynet.cms.gov/inpatient/measures/readmission/methodology.

Two or more high-risk prescriptions for medications in the same medication class. This measure approximates the Healthcare Effectiveness Data and Information Set (HEDIS®) High Risk Medications in the Elderly measure that is included in the Quality Payment Program (QPP). We used the 2021 specifications (HEDIS; available at Quality ID #238 (NQF 0022): Use of High-Risk Medications in Older Adults [https://qpp.cms.gov/docs/QPP_quality_measure_specifications/CQM-Measures/2021_Measure_238_MIPSCQM.pdf]). It is defined as the percentage of beneficiaries age 65 and older who received two or more medications with a high risk designation within the same class. A lower rate indicates better performance.

The denominator includes beneficiaries who were: at least 65 years old at the beginning of the measurement year and continuously enrolled in Medicare Parts A, B, and D during the measurement year. The denominator excludes those who used hospice services in the measurement year. The numerator is based on year-specific value sets from the National Committee for Quality Assurance (NCQA) that contain National Drug Codes (NDC) that map the medication classes. Examples of "high-risk" medication classes include antispasmodics, antithrombotics, and non-benzodiazepine hypnotics. Table 5.C.16 presents a list of all the medication classes, which can also be found in the QPP documentation. We did not require a clinician encounter for inclusion in the numerator. To align with the HEDIS specification, we did not report the rate of receipt of one high-risk medication as described in the QPP documentation. (The value sets are available at: https://www.ncqa.org/hedis/measures/.)

Table 5.C.16. High-risk medication drug classes

Anticholinergics, first-generation antihistamines	Endocrine system, sulfonylureas, long-duration
Anticholinergics, anti-Parkinson agents	Endocrine system, other
Antispasmodics	Pain medications, skeletal muscle relaxants
Antithrombotics	Pain medications, other
Cardiovascular, alpha agonists, central	Anti-infectives, other[a]
Cardiovascular, other	Nonbenzodiazepine hypnotics[b]
Central nervous system, antidepressants	Alpha agonists, central[c]
Central nervous system, barbiturates	Cardiovascular, other[c]
Central nervous system, vasodilators	Tertiary TCAs (as single agent or as part of combination products)[c]
Central nervous system, other	
Endocrine system, estrogens with or without progestins; includes only oral and topical patch products	

Note: Unless otherwise noted, medications with any dose or duration in these classes are considered high risk.

[a] Medication with days-supply criteria.

[b] Medication with days-supply criteria prior to 2020, but considered high risk with any dose in 2020 and 2021. See this change in the NCQA documentation available at https://www.ncqa.org/wp-content/uploads/2020/07/20200716_Summary_Table_of_Measures_Product_Line_and_Changes_UPDATED.pdf.

[c] Medication with average daily dose criteria.

TCAs = tricyclic antidepressants.

Long-term opioid use and **Potential opioid overuse**. Following the specifications for the denominator for the electronic Clinical Quality Measure (eCQM) 460 (eCQI Resource Center, available at https://ecqi.healthit.gov/sites/default/files/ecqm/measures/CMS460v2.html), we defined long-term opioid use as having an opioid supply of 90 days or more in one year with no more than a 7-day gap between

prescriptions. Using the same specifications, potential opioid overuse was defined as the use of opioids at a daily dosage of 90 morphine milligram equivalents (MMEs) or more among long-term opioid users.[42]

A beneficiary had to (1) be continuously enrolled in Medicare Parts A, B, and D throughout each calendar year or until death, and (2) have at least one opioid prescription during the measurement year (that is, had to have some opioid use) to be included in these measures. Because eCQM 460 does not list national drug codes (NDCs), to identify beneficiaries who used opioid therapy, we relied on NDCs for opioid therapy from the Medication List Directory value sets for the HEDIS® measure of high-dosage opioid use (NCQA 2020, 2021, and 2022 available at https://store.ncqa.org/hedis-my-2020-medication-list-directory.html, https://store.ncqa.org/hedis-my-2021-medication-list-directory.html, and https://store.ncqa.org/hedis-my-2022-medication-list-directory.html). We used the CDC Opioid NDC and the Oral MME Conversion File (CDC, available at https://www.cdc.gov/opioids/data-resources/index.html) to calculate daily MME for beneficiaries on opioid therapy.

We excluded beneficiaries for whom opioid use is appropriate: those with a diagnosis of cancer during or one year before the measurement year, those with a diagnosis of sickle cell disease, and those with any hospice use during the measurement year. To identify diagnoses for exclusion criteria, we used ICD-10 codes from eCQM specifications. Even though potential opioid overuse excludes most of the beneficiaries for whom such use is appropriate (those with cancer or sickle cell disease, and those who use hospice), it does not take *all* appropriate use into account, such as use of opioids in non-hospice palliative care.

Number of low-value services per 1,000 beneficiaries per year. This measure is the annualized total number of services that "provide little to no benefit to patients, have potential to cause harm, incur unnecessary costs to patients, or waste limited healthcare resources" (Maratt et al. 2019), per 1,000 beneficiaries. Building on the work of Schwartz and colleagues (2014, 2015), we use 31 services including 28 services used in the original Schwartz measure and 3 services identified as part of our update (Fleming et al. 2022).[43] Table 5.C.17 lists the 31 services and their definitions. To be included in this measure, a beneficiary must be continuously enrolled in Medicare FFS Parts A and B, be alive at the end of the measurement periods, have a record in the MBSF, and meet service-specific criteria during the measurement periods.[44] Because three of the low-value services are identified using a one-year look-forward period to determine whether a service was or was not low value, this measure is only available for the years 2016–2020.

[42] The main difference between our potential opioid overuse measure and eCQM 460 is that our measure relies on Part D claims data, whereas eCQM 460 relies on electronic health record (EHR) data.

[43] The three services identified in our update are: (1) laminectomy or spinal fusion for individuals without clear indications of radicular pain or herniated disc, (2) electromyography (EMG) for low back pain among individuals without diagnoses of leg pain or sciatica, and (3) magnetic resonance imaging (MRI) of the peripheral joints to monitor rheumatoid arthritis.

[44] For example, the service of prostate-specific antigen (PSA) tests requires beneficiaries to be 75 years of age and older with no history of prostate cancer.

Table 5.C.17. Low-value services

Service	Description
Cancer Screening	
Cancer screening for patients with CKD receiving dialysis	Screening for cancer of the breast, cervix, colon, or prostate for patients age 75 and over with CKD receiving dialysis services
Cervical cancer screening for women age 65 and over	Screening Papanicolaou test for women age 65 and over with no personal history of cancer or dysplasia noted in claim or in prior claims, and no diagnoses of other female genital cancers, abnormal Papanicolaou findings, or human papillomavirus positivity in prior claims
Colorectal cancer screening for adults over age 85	Colorectal cancer screening (colonoscopy, sigmoidoscopy, barium enema, or fecal occult blood testing) for patients age 86 or over with no history of colon cancer
PSA testing for men age 75 and over	PSA testing for patients age 75 and over with no history of prostate cancer
Diagnostic and Preventive Testing	
Bone mineral density testing at frequent intervals	Bone mineral density test within two years of a prior bone mineral density test for patients with established osteoporosis diagnosis
Homocysteine testing in cardiovascular disease	Homocysteine testing with no diagnoses of folate or B12 deficiencies in the claim and no folate or B12 testing in prior claims
Hypercoagulability testing for patients with deep vein thrombosis (DVT)	Lab tests for hypercoagulable states within 30 days following diagnosis of lower extremity DVT or pulmonary embolism; no prior evidence of recurrent thrombosis, defined by diagnosis of DVT or pulmonary embolism more than 90 days before the testing claim
PTH measurement for patients with stage 1-3 CKD	PTH measurement for patients with CKD and no dialysis services before PTH testing or within 30 days following testing, as well as no hypercalcemia diagnosis during the year
Total or free T3 level testing for patients with hypothyroidism	Total or free T3 measurement in a patient with hypothyroidism diagnosis during the year
1,25-dihydroxyvitamin D testing in the absence of hypercalcemia or decreased kidney function	Calcitriol testing for patients without hypercalcemia, secondary hyperparathyroidism of renal origin, or conditions related to non-PTH-mediated hypercalcemia noted in claim (sarcoidosis, tuberculosis, selected neoplasms), and without a history of CKD; no diagnosis of hypercalcemia in the past 30 days
EMG for low back pain[a]	Individuals without diagnoses of leg pain or sciatica who receive EMG studies
Preoperative Testing	
Preoperative echocardiography	Echocardiogram not associated with inpatient or emergency care and occurring within 30 days before a low or intermediate risk noncardiothoracic surgical procedure
Preoperative PFT	PFT not associated with inpatient or emergency care and occurring within 30 days before a low or intermediate risk surgical procedure
Routine preoperative stress tests	Stress electrocardiogram, echocardiogram, nuclear medicine imaging, cardiac MRI, or CT angiography, not associated with inpatient or emergency care and occurring within 30 days before a low or intermediate risk surgical procedure
Imaging	
Computed tomography (CT) of the sinuses for uncomplicated acute rhinosinusitis	Maxillofacial CT study with a diagnosis of sinusitis and no complications of sinusitis, immune deficiencies, nasal polyps, or head/face trauma noted in claim and no sinusitis diagnosis between 30 and 365 days before imaging
Head imaging in the evaluation of syncope	CT or MR imaging of the head with a diagnosis of syncope and no diagnoses in claim warranting imaging
Head imaging for uncomplicated headache	Brain CT or MR imaging with non-post-traumatic, non-thunderclap headache diagnosis, and no diagnoses in claim warranting imaging
Electroencephalogram (EEG) for headaches	EEG with headache diagnosis in claim, and no epilepsy or convulsions noted in current or prior claims
Back imaging for patients with nonspecific low back pain	Back imaging with a diagnosis of lower back pain occurring within six weeks of initial back pain diagnosis and with no indication of radiculopathy or other diagnoses in claim warranting imaging

Table 5.C.17. ***(continued)***

Service	Description
Screening for carotid artery disease in asymptomatic adults	Carotid imaging not associated with inpatient or emergency care for patients without a history of stroke or TIA, and without a diagnosis of stroke, TIA, or focal neurological symptoms in claim
Screening for carotid artery disease for syncope	Carotid imaging with syncope diagnosis for patients without a history of stroke or TIA, and without a diagnosis of stroke, TIA, or focal neurological symptoms in claim
Imaging for diagnosis of plantar fasciitis	Radiographic or magnetic resonance imaging (MRI) with diagnosis of plantar fasciitis occurring within two weeks of initial foot pain diagnosis
MRI for rheumatoid arthritis (RA)[a]	Individuals who receive an MRI of the peripheral joints following a diagnosis of RA
Cardiovascular Testing and Procedures	
Stress testing for stable coronary disease	Stress testing not associated with inpatient or emergency care for patients with an established diagnosis of acute myocardial infarction (6 months or more before testing)
Percutaneous coronary intervention with balloon angioplasty or stent placement for stable coronary disease	Coronary stent placement or balloon angioplasty, not associated with an ER visit, or patients with an established diagnosis of acute myocardial infarction (greater than or equal to 6 months before testing)
Renal artery angioplasty or stenting	Renal/visceral angioplasty or stent placement with a diagnosis of renal atherosclerosis or renovascular hypertension noted in procedure claim
Carotid endarterectomy for asymptomatic patients	Carotid endarterectomy, not associated with an emergency room visit, for female patients without a history of stroke or TIA and without stroke, TIA, or focal neurological symptoms noted in claim
Pulmonary artery catheterization in the ICU	Pulmonary artery catheterization for monitoring purposes during an inpatient stay that involved an ICU and nonsurgical MS-DRG; claim contains no diagnoses indicating pulmonary hypertension, cardiac tamponade, or preoperative assessment
Other Invasive Procedures	
Vertebroplasty or kyphoplasty for osteoporotic vertebral fractures	Vertebroplasty/kyphoplasty for vertebral fracture, with no bone cancers, myeloma, or hemangioma noted in procedure claim
Spinal injection for low back pain	Outpatient epidural, facet, or trigger point injections for lower back pain, excluding etanercept; no radiculopathy diagnoses in the claim
Laminectomy or spinal fusion[a]	Individuals without clear indications of radicular pain or of herniated disc who receive a laminectomy and/or spinal fusion

[a] Indicates additional measures from Fleming et al. (2022).

CKD = chronic kidney disease; CT = computed tomography; DVT = deep vein thrombosis; EEG = electroencephalogram; EMG = electromyography; ICU = intensive care unit; PFT = pulmonary function testing; PTH = parathyroid hormone; RA = rheumatoid arthritis.

H. Mortality

We constructed annual measures of mortality and days a beneficiary was alive for Medicare FFS beneficiaries attributed in the first quarter of baseline or the first quarter of the intervention:

- 12-month mortality in the year before the start of CPC+: percentage who died within 12 months (by the end of the Baseline year)
- 12-month mortality: percentage who died within 12 months (by the end of PY 1)
- 24-month mortality: percentage who died within 24 months (by the end of PY 2)
- 36-month mortality: percentage who died within 36 months (by the end of PY 3)
- 48-month mortality: percentage who died within 48 months (by the end of PY 4)
- 60-month mortality: percentage who died within 60 months (by the end of PY 5)
- 12-month survival in the year before the start of CPC+: fraction of days alive across 12 months (by the end of the Baseline year)
- 12-month survival: fraction of days alive across 12 months (by the end of PY 1)
- 24-month survival: fraction of days alive across 24 months (by the end of PY 2)
- 36-month survival: fraction of days alive across 36 months (by the end of PY 3)
- 48-month survival: fraction of days alive across 48 months (by the end of PY 4)
- 60-month survival: fraction of days alive across 48 months (by the end of PY 5)

5.C.2. Non-outcome claims-based measures

We quantify how participation in other initiatives differs between CPC+ and comparison practices and how this participation shifted from the baseline period to the first three program years of CPC+ for each group (Appendix 5.E). We discuss two broad types of CMS initiatives below: care management services and behavioral integration services.

Receipt of chronic care management, transitional care management, or other care management services. We used these three measures to examine the extent of receipt of each type of care management services as well as any care management services during the year by beneficiaries assigned to CPC+ and comparison practices. We identified beneficiaries with a claim in the carrier or outpatient file with one of the procedure codes in Table 5.C.18 as having received one of these management services. Comparable to the ambulatory visit specifications, we did not include add-on services in our algorithm. The CPT Editorial Panel instituted several procedure code updates during our analytic time period, so our specifications were updated to reflect codes as they were added, deleted, or replaced. We included new procedure codes as they were implemented or updated them when they were replaced. In 2021, we added HCPCS Codes G2214: Psychiatric Collaborative Care Management – 1st 30 minutes. The last column of Table 5.C.18 shows the time period during which each procedure code was used. Although CPC+ practices cannot bill chronic care management services for attributed Medicare beneficiaries, we expect to observe a small proportion of CPC+ beneficiaries with such claims in our analysis sample based on intent-to-treat assignment rules, under which we retain beneficiaries even if they are no longer attributed to a CPC+ practice.

Receipt of general behavioral health integration and psychiatric collaborative care management. In January 2017, CMS introduced FFS Medicare Part B billing codes for Psychiatric Collaborative Care Management (CoCM) and General Behavioral Health Integration (BHI) (CMS 2019a). CoCM enhances primary care through the addition of behavioral health care managers and psychiatric consultation, whereas BHI supports various integration models and staffing configurations. We created three indicators at the beneficiary level for receipt of behavioral health care management services during the intervention years: (1) BHI, (2) psychiatric CoCM, and (3) psychiatric collaborative care model at an FQHC or RHC. These indicators are subsets of the existing chronic and other care management categories that we describe above and note in Table 5.C.18.

Table 5.C.18. Procedure codes for chronic care management, transitional care management, and other care management services

	CPT/HCPCS code	Description	Time period during which procedure code is included in measures
Chronic care management	99490	Chronic care management (20 minutes of clinical staff time)	2016–2021
	99491	Chronic care management (30 minutes of clinical staff time)	2019–2021
	99487	Complex chronic care management (60 minutes of clinical staff time)	2017–2021
	99484[a]	General behavioral health integration care management	2018–2021
	G0506	Chronic care management care planning	2017–2021
	G0507[a]	Care management services for behavioral health conditions	2017 (deleted in 2018 and replaced with 99484)
	99358	Prolonged (<75 minutes) of non-face-to-face E&M service before and/or after direct patient care	2016–2021
Transitional care management	99495	Transitional care management for patients discharged to community from an inpatient setting; moderate complexity of medical decision making	2016–2021
	99496	Transitional care management for patients discharged to community from an inpatient setting; high complexity of medical decision making	2016–2021
Other care management	G0181	Home health supervision of at least 30 minutes	2016–2021
	G0182	Hospice health supervision of at least 30 minutes	2016–2021
	G0502[b]	Initial psychiatric collaborative care management, first 70 minutes	2017 (Deleted in 2018 and replaced with 99492)
	G0503[b]	Subsequent psychiatric collaborative care management, first 60 minutes	2017 (Deleted in 2018 and replaced with 99493)
	G0504[b]	Initial or subsequent psychiatric collaborative care management, additional 30 minutes	2017 (Deleted in 2018 and replaced with 99494)
	G0505	Cognition and functional assessment	2017 (Deleted in 2018 and replaced with 99483)
	G0511	General care management at an FQHC or RHC	2018–2021
	G0512[c]	Psychiatric collaborative care model at an FQHC or RHC	2018–2021
	G2064	Principal care management (physicians and non-physicians)-covers services for patients with only one complex chronic condition that requires management by a specialist	2021

	CPT/HCPCS code	Description	Time period during which procedure code is included in measures
	G2065	Principal care management (clinical staff)- covers services for patients with only one complex chronic condition that requires management by a specialist	2021
	99483	Cognitive assessment	2018–2021
	99492[b]	Initial psychiatric collaborative care management	2018–2021
	99493, 99494[b]	Subsequent psychiatric collaborative care management	2018–2021
	G2214[b]	Psychiatric Collaborative Care Management – 1st 30 minutes	2021
	99497	Advance care planning	2016–2021

Sources: American Medical Association. "CPT, Professional Edition." 2016–2021; American Medical Association. "HCPCS Level II, Professional Edition." 2016–2021.

Note: CPT Codes 99489 (Additional 30 minutes of clinical staff time for chronic care management) and 99359 (Additional 30 minutes of prolonged non-face-to-face E&M service before and/or after direct patient care) were used to identify CCM services for our first annual report but were not used to identify CCM services in subsequent reports.

[a] General Behavioral Health Integration (BHI)

[b] Psychiatric Collaborative Care Management (CoCM)

[c] Psychiatric collaborative care model at an FQHC or RHC

CCM = chronic care management; CPT = Current Procedural Terminology; E&M = Evaluation and Management; FQHC = Federally Qualified Health Center; HCPCS = Health Care Common Procedure Coding System; OCM = other care management; RHC = Rural Health Center; TCM = transitional care management.

5.C.3. Claims-based control and subgroup variables

In this section, we discuss the construction of claims-based control variables we used in our regression analysis that all center on beneficiary health and chronic conditions.

Three beneficiary-level claims-based control variables were derived from the hierarchical condition category (HCC) software: (1) an HCC score, which is a measure of risk for subsequent expenditures; (2) an indicator for "new enrollees"; and (3) indicators for 21 chronic condition categories. We also created an indicator for Alzheimer's disease or dementia based on the Chronic Conditions Warehouse (CCW) algorithm. We describe these measures below.

Hierarchical condition category score. We controlled for HCC score in our regressions to account for variation in beneficiaries' health status, or their level of risk for Medicare spending (Pope et al. 2004, 2011). We controlled for the baseline HCC score (calculated using 2015 claims for beneficiaries attributed to practices that started in 2017) for observations in the baseline period. To avoid endogeneity issues, we controlled for the score at the start of the intervention (calculated using 2016 claims for beneficiaries attributed to practices that started in 2017) for observations during the entire intervention period (i.e., we did not update the HCC score during the intervention period with claims data drawn from the intervention period). We also include a binary control variable in our regression analysis that indicates whether the HCC score was calculated using only demographic information.[45]

We calculated both the baseline and intervention period HCC scores using CMS's HCC score software and algorithm, based on information from Medicare claims and enrollment data. We deviated from the

[45] HCC scores are calculated on the basis of demographic characteristics only when claims data are not observed for a beneficiary and may not reflect the actual risk of the beneficiary. This situation generally happens when the beneficiary is new to Medicare FFS.

exact approach CMS uses in a few ways to adapt the CMS algorithm for the purpose of the impact analysis. For instance, to avoid endogeneity concerns, we used information on dual status, long-term institutionalization (LTI), and ESRD status from the prior year instead of the year for which the HCC score was being calculated. Also, we adopted a more nuanced approach to assigning the new enrollee versus the community score to beneficiaries with less than 12 months of FFS enrollment during the base year, as described in Step 5 below.

Specifically, we used the following approach:

1. To calculate HCC scores, we continued to use Version 22 2017 HCC model software,[46] which has greater predictive accuracy than earlier versions. We also used the Version 21 2017 ESRD model software for beneficiaries with ESRD.
2. To calculate HCC scores, we used a 12-month lookback for Medicare claims to obtain diagnosis information. For instance, to calculate the 2017 HCC score, we used Medicare claims during 2016. For beneficiaries that are newly attributed after 2017, we still use their 2016 Medicare claims (if they exist) to calculate their 2017 HCC score.
3. The HCC algorithm also uses information on demographics, reason for Medicare eligibility, new enrollee status, dual eligibility status (with the latest version of the model distinguishing between beneficiaries who have full versus partial dual eligibility status), long-term nursing home care, kidney transplant, and dialysis status. To estimate and assign HCC scores for any year, we used information on these attributes from the prior year, with the exception of demographics and reason for Medicare eligibility, which were from the current year. For example, to calculate the 2017 HCC score, we used the following beneficiary information:
   - Demographics from 2017
   - Medicare eligibility (eligible due to age or disability) from 2017
   - New enrollee status from 2016 (a beneficiary with less than six months of Medicare FFS enrollment during the year was flagged as a new enrollee)
   - Dual eligibility status (full, partial, or nondual) during the last three months of 2016
   - ESRD status during the last three months of 2016
   - LTI status during a 120-day period ending on December 31, 2016
   - The number of months since a kidney transplant, looking back from January 1, 2017
   - Whether the transplant was successful or the beneficiary was on dialysis
4. The HCC algorithm estimates the following separate models: (1) ESRD (further differentiating by dialysis status and time since kidney transplant), (2) LTI, (3) community (further differentiating by dual status and aged versus disabled status), and (4) new enrollee. These models include different covariates and interaction terms, and therefore lead to multiple values of the HCC scores for each beneficiary. For instance, the new enrollee model is estimated with covariates only for demographics and Medicare eligibility information, without any covariates for claims-based diagnoses. Thus, for the 2017 HCC score, a beneficiary would have multiple values with one score from each model.

[46] We have incorporated the 2018–2021 ICD-10 codes into the Version 22 2017 software.

5. After estimating the four HCC models, we selected one HCC score for each beneficiary, following CMS's approach to determine which model's score was appropriate for the beneficiary. For example, we assigned a specific value of the 2017 HCC score to a beneficiary, by progressively checking the criteria in the following order:
   - We assigned the value of the ESRD score to a beneficiary for the 2017 HCC score if the beneficiary had ESRD anytime during the last three months of 2016 (the ESRD score could further vary or could come from a different ESRD submodel, depending on length of time since a successful kidney transplant, dialysis status, new enrollee status, and age).
     - We rescaled the risk scores for ESRD and post-kidney transplant beneficiaries to account for the fact that their average costs differ from the average costs for the overall FFS population. For ESRD beneficiaries on dialysis, their 2016 and 2017 HCC scores were multiplied by factors of 8.146 and 8.227, respectively. For beneficiaries with functioning grafts, multiplication factors were 0.866 (2016 HCC score) and 0.875 (2017 score).[47]
   - If a beneficiary did not have ESRD and met the criteria for LTI during the 120-day period ending on December 31, 2016, we assigned the value of the institutional or LTI score for 2017.
   - If a beneficiary did not meet the criteria for either the ESRD or LTI score, and:
     - Had less than six months of Medicare FFS enrollment during 2016, we assigned the new enrollee score for 2017. (Note that this approach is used for baseline scores as well.)
     - Had 10 or more months of Medicare FFS enrollment during 2016, we assigned the community score for 2017. The community score varied or was obtained from a different submodel, depending on dual status (full, partial, or nondual) during the last three months of 2016, and aged versus disabled status.
     - Had six to nine months of Medicare FFS enrollment during 2016, we again assigned the community score for 2017 (varying as above by dual and aged or disabled status) but adjusted that score upward or inflated it by 25 percent. We used this approach to account for missing information on Medicare claims for three to six months in 2016, and therefore, the limited information on diagnoses available for such beneficiaries.
6. Finally, we used CMS's official normalization factors for 2016 and 2017 HCC scores to calculate a normalized risk score for each beneficiary. Specifically, the normalized risk score for 2016 (or 2017) is equal to the raw 2016 (or 2017) risk score, calculated using the approach laid out above, divided by the normalization factor for that year. The normalization factors account for changes in coding practice as well as in population demographics between the year an HCC model was calibrated and the year for which we calculated the HCC score.

[47] The resource for the ESRD rescaling factors is the CCW Geographic Variation Database (GVDB) V5 manual.

Indicator for whether a beneficiary was assigned a new enrollee score. Our regressions also controlled for whether a beneficiary was assigned a new enrollee score in the baseline or intervention period. The other types of scores (community, LTI, ESRD, etc.) are based on the beneficiary's actual claims history, but the new enrollee score (which is assigned to beneficiaries with less than six months of FFS eligibility during the lookback period) is only a proxy for the beneficiary's actual risk, because it is based only on the beneficiary's demographic characteristics and reason for Medicare entitlement. A beneficiary that is first attributed after 2017 and is assigned a new enrollee score (based on having less than six months of claims or no claims in 2016) will retain that same score throughout the entire intervention period. The scores are not updated, because they could be affected by the care that the beneficiary receives during the intervention.

Chronic condition indicators based on individual or combined HCCs. In addition to HCC scores, our regressions also controlled for HCCs. The HCC models produce the HCCs as part of generating the HCC score by using diagnosis information in Medicare claims (Pope et al. 2004, 2011). The models produce a total of 87 HCCs (79 from the V22 HCC model and an additional 8 from the ESRD model). Based on investigations for our first annual report, we had identified 21 HCCs (Table 5.C.19) to include as control variables to adjust for chronic conditions in our regressions, in three steps outlined below. We continued to use the same HCCs in this report, creating baseline and intervention period versions. The baseline measures are based on diagnoses in the prior year or the pre-baseline year (2015). The measures used during the intervention period (Years 1 through 3) are based on diagnoses in the baseline year (2016). Note that a beneficiary will never have a condition in the intervention period if the beneficiary has no claims in 2016. The indicator for the new enrollee score enables us to distinguish between true zeroes on these conditions (beneficiaries that had claims, but did not have the condition) versus those that do not show up as having the condition because they did not have claims in 2016.

Step 1. We narrowed the pool to 38 HCCs that met at least one of the following criteria:

- Had a relatively high prevalence among beneficiaries in our sample (4 percent and above).
- Had higher than average relative factors (greater than or equal to 1) from the HCC models, implying that they were important predictors of Medicare expenditures.
- Showed a noticeable change in prevalence rates between the baseline year (2016) and the follow-up year (2017), among beneficiaries in the yearly samples (greater than or equal to 0.4 percentage points in the CPC+ group or the comparison group).
- Showed a noticeable difference in prevalence rates between CPC+ and comparison beneficiaries in the sample (greater than or equal to 0.2 percentage points).

Step 2. We ran difference-in-differences regressions for Medicare expenditures without fees, using one year of baseline period data and one year of follow-up period data, and including all 38 HCCs, separately for Track 1 and Track 2 practices.

Step 3. Based on the magnitude and significance of the coefficient estimate for each HCC in these regressions, and their overall prevalence in our sample, we selected 21 categories as regression controls (Table 5.C.19). Eleven of these HCCs were individual HCCs denoting a specific condition, and the 10 others were combinations of one or more HCCs. We combined certain HCCs with high or statistically significant coefficient estimates if their individual rates of prevalence were low and they belonged to the same broad family of conditions.

Table 5.C.19. List of hierarchical condition categories used as chronic condition controls

Hierarchical condition category	Description
HCC 8	Metastatic Cancer and Acute Leukemia
HCC 18	Diabetes with Chronic Complications
HCC 21	Protein-Calorie Malnutrition
HCC 22	Morbid Obesity
HCC 23	Other Significant Endocrine and Metabolic Disorders
HCC 85	Congestive Heart Failure
HCC 96	Specified Heart Arrhythmias
HCC 106	Atherosclerosis of the Extremities with Ulceration or Gangrene
HCC 111	Chronic Obstructive Pulmonary Disease
HCC 173	Traumatic Amputations and Complications
HCC 186	Major Organ Transplant or Replacement Status
HCC 40 or 47	Rheumatoid Arthritis and Inflammatory Connective Tissue Disease or Disorders of Immunity
HCC 46 or 48	Severe Hematological Disorders, or Coagulation Defects and Other Specified Hematological Disorders
HCC 54 or 55	Drug/Alcohol Psychosis or Dependence
HCC 57 or 58	Schizophrenia or Major Depressive, Bipolar, and Paranoid Disorders
HCC 70 or 71	Quadriplegia or Paraplegia
HCC 80 or 82	Coma, Brain Compression/Anoxic Damage or Respirator Dependence/Tracheostomy Status
HCC 86, 87, or 88	Acute Myocardial Infarction, Unstable Angina and Other Acute Ischemic Heart Disease, or Angina Pectoris
HCC 99 or 100	Cerebral Hemorrhage, or Ischemic or Unspecified Stroke
HCC 107 or 108	Vascular Disease, with Complications
HCC 157 or 158	Pressure Ulcer of Skin with Necrosis Through to Muscle, Tendon, or Bone; or of Skin with Full Thickness Skin Loss

Source: Centers for Medicare and Medicaid Services. "CMS-HCC Risk Adjustment Model." 2017–2018. Available at https://www.cms.gov/Medicare/Health-Plans/MedicareAdvtgSpecRateStats/Risk-Adjustors.

Indicator for presence of Alzheimer's disease or dementia based on the CCW algorithm. Similar to the HCCs described above, we constructed a CCW indicator for Alzheimer's disease or dementia to adjust for this condition in our regressions. (This indicator is also used to identify high-risk beneficiaries in risk Tier 5, as described in Chapter 5 in Peikes et al. 2021b.) We used this CCW indicator instead of HCCs for Alzheimer's disease and dementia from the HCC model to ensure consistency with CMS's approach for identifying high-risk, Tier 5 beneficiaries in Track 2 of CPC+. We created annual indicators based on the CCW algorithm, which uses a three-year lookback period to identify these diagnoses. For example, our baseline (2016) indicator used claims from January 1, 2013, through December 31, 2015, and our indicator for Alzheimer's and dementia at the start of the intervention period (2017) used claims from January 1, 2014, through December 31, 2016.

The CCW algorithm for defining this indicator requires a diagnosis code from Table 5.C.20 in any position on at least one inpatient, skilled nursing facility, home health, outpatient, or carrier claim during the three-year lookback period.

Table 5.C.20. Diagnosis codes used to identify Alzheimer's disease or dementia

ICD-9 diagnosis codes	ICD-10 diagnosis codes
331.0, 331.11, 331.19, 331.2, 331.7, 290.0, 290.10, 290.11, 290.12, 290.13, 290.20, 290.21, 290.3, 290.40, 290.41, 290.42, 290.43, 294.0, 294.10, 294.11, 294.20, 294.21, 294.8, 797	F01.50, F01.51, F02.80, F02.81, F03.90, F03.91, F04, G13.8, F05, F06.1, F06.8, G30.0, G30.1, G30.8, G30.9, G31.1, G31.2, G31.01, G31.09, G94, R41.81, R54

Source: Centers for Medicare & Medicaid Services. "Chronic Conditions Data Warehouse (CCW)." 2016–2021. Available at https://www2.ccwdata.org/web/guest/condition-categories.

Indicator for anxiety or depression. For the final report, we updated the beneficiary-level mental health subgroup definition to focus on beneficiaries with anxiety or depression. Previously, we had used the hierarchical condition category (HCC)-based behavioral health subgroup defined as beneficiaries with schizophrenia (HCC 57), or major depressive, bipolar, and paranoid disorders (HCC 58). However, this subgroup definition is likely to be limited because it includes beneficiaries with severe behavioral health conditions and it excludes beneficiaries with anxiety and other forms of depression. The behavioral health integration models[48] that CPC+ practices implement have been primarily tested among primary care populations with depression and anxiety rather than on those with more severe or chronic behavioral health conditions.

For our updated definition, we created anxiety and depression indicators for the baseline period (2016) using claims directly before the period (2015), and indicators for the intervention period using claims directly before the intervention period (2016 claims). We flagged diagnoses on the acute inpatient (including CAHs), outpatient, SNF, carrier (excluding DME), and non-acute psychiatric inpatient claims. The diagnoses used to identify anxiety and depression are presented in Table 5.C.21.

Table 5.C.21. Diagnosis codes used to identify anxiety or depression

Category	ICD-9 diagnosis codes	ICD-10 diagnosis codes
Anxiety	300.0, 300.2, 300.3, 308, 309.21, 309.81, 293.84	F40–F43, F93.0, F06.4
Depression	296.2, 296.3, 300.4, 311	F32, F33, F34.1

Sources: American Medical Association. "ICD-9-CM: The Complete Official Codebook." 2015; American Medical Association. "ICD-10-CM: The Complete Official Codebook." 2016.

As before with the HCC approach, we excluded diagnoses on the carrier and outpatient files that were associated with non-covered facilities (free-standing Ambulatory Surgical Centers (ASCs), home health care, and free-standing renal dialysis facilities), laboratory and imaging services (except for those marked for inclusion), and HCC-excluded providers (e.g., diagnostic radiology, clinical laboratory, multispecialty clinic, dietician). We further restricted the subgroup to those beneficiaries who had at least one inpatient or two outpatient claims with the required diagnoses in the baseline period, with anxiety and depression identified separately. This protected us from including beneficiaries with transient or potentially misdiagnosed conditions. We then kept the beneficiaries who had claims that met the above criteria and had either an anxiety flag, a depression flag, or both. Although diagnosis codes in Medicare claims and other administrative data tend to under-identify depression relative to screening tools and medical records (Townsend et al. 2012; Noyes et al. 2011), studies have also found that requiring at least two depression-related visits to a health care provider improves the specificity of identifying cases (Solberg et al. 2006; Kerr et al. 2000). Our final exclusion was to drop beneficiaries who had comorbid dementia or Alzheimer's disease diagnoses, as identified by the CCW algorithm described above. Beneficiaries with comorbid dementia or Alzheimer's disease are not the target of CPC+ integration activities and likely require more intensive supports.

[48] Additional information is presented at: https://innovation.cms.gov/Files/x/cpcplus-bhinteg-options.pdf.

5.C.4. Non-claims-based control variables

For beneficiary-level analyses, we controlled for beneficiaries' demographics (age, race, and gender) and original reason for Medicare eligibility (age, disability, or ESRD) in our regression models, based on information in the Medicare enrollment database. We calculated age as of January 1 of the baseline year for the baseline observations (2016), and as of January 1 of the first intervention year (2017) for observations in the intervention period. We describe the exact age and race categories used in our regressions in Appendix 5.C.

We also controlled for dual eligibility status, based on information obtained from the Master Beneficiary Summary File (MBSF). Specifically, we used the DUAL_STATUS_CD variable in the MBSF during the last three months of the pre-baseline (2015) and baseline (2016) years to define dual status for the baseline and intervention periods, respectively. We flagged a beneficiary as dually eligible if this variable indicated either full or partial dually eligible status during any of those three months.[49] For beneficiaries who enrolled in Medicare after the three months prior to the measurement period (i.e., the last three months of 2015 or the last three months of 2016), we assigned the non-dual status for the corresponding measurement period by default, because they did not have a dual status in the MBSF before their enrollment. For example, if a beneficiary enrolled in Medicare in 2016, then we assigned the non-dual status for the baseline period, because the beneficiary did not have a dual status in the MBSF during the last three months of 2015. Similarly, if a beneficiary enrolled in Medicare in 2018, then we assigned the non-dual status for all intervention periods, because the beneficiary did not have a dual status in the MBSF during the last three months of 2016. Consistent with our approach for other covariates, we do not update the dual status during a measurement period, because it could be affected by the care that the beneficiary receives during the intervention.

[49] We used dual eligibility status in the three months *prior to the measurement period* (baseline or first intervention year) as a control variable to avoid endogeneity concerns with using concurrent values of time-varying beneficiary characteristics. Using the *last three months* before the start of the measurement period for outcomes gives us the closest approximation to dual status during the measurement period. This approach differs from CMS's dual status specification for payment purposes, in which concurrent month-by-month dual status is used to determine the appropriate risk score in the month.

5.D. Implications of COVID-19 for the CPC+ impact evaluation

The coronavirus 2019 (COVID-19) pandemic could introduce bias into our impact estimates for the CPC+ evaluation if COVID-19 differentially affected outcomes for CPC+ and comparison regions. In this Appendix, we evaluate the likelihood that COVID-19 biased impact estimates during Program Year (PY) 4 and PY 5. We first introduce the motivation and research questions for the analysis (Section 5.D.1). We then report the effects of COVID-19 in CPC+ and comparison regions, including the direct effects of COVID-19 on the prevalence of diagnoses and excess deaths (Section 5.D.2), as well as the total effects of COVID-19 on health care utilization and Medicare expenditures (Section 5.D.3). We next describe how we used regional COVID-19-related controls to account for differences in the CPC+ and comparison regions due to COVID-19 (Section 5.D.4). Finally, we discuss the key findings from the analysis and their implications for the CPC+ impact evaluation (Section 5.D.5). We include additional results and methodological details for this analysis in supplemental Sections 1–8.

5.D.1. Introduction

The CPC+ impact evaluation relies on comparison practices selected from "external" regions—here defined as states or contiguous counties that did not have any CPC+ practices. These regions could have experienced effects of the COVID-19 pandemic that were different than those experienced by CPC+ regions. The timing and magnitude of the pandemic differed considerably by region, particularly early in the pandemic (Oster et al. 2020), leading to concerns that COVID-19 could introduce bias into the CPC+ impact evaluation.

In this Appendix, we examine whether COVID-19 differentially affected key outcomes in CPC+ versus comparison regions. This helps us evaluate the likelihood that COVID-19 biased impact estimates for the CPC+ evaluation in PY 4 and PY 5. Our analytical approach for the CPC+ impact evaluation for the fifth annual report (AR5) was based, in part, on the findings from this analysis. In a series of similar checks in the fourth annual report (AR4), we found there were small regional differences in the effects of COVID-19 between CPC+ and comparison regions in 2020. Further, we found that adding COVID-19-related controls to our regression models often reduced differences by more than one-half for the three outcomes we studied (expenditures for Medicare Part A and B services without enhanced payments, all-cause acute hospitalizations, and outpatient emergency department [ED] visits). However, the regional differences we observed in 2020 due to COVID-19 in AR4 analyses could differ from those in 2021: the second year of the pandemic was marked by the availability of vaccines and at-home testing; waves of regional hotspots; and a greater return towards business as usual.

What's new this year?

1. Additional year of data (2021)
2. Stratified results by Medicare Shared Savings Program (SSP) status within each track
3. Additional acute care outcomes, potentially influenced by COVID-19
   - Expenditures for acute hospitalizations
   - Expenditures for outpatient ED visits
   - Urgent care center (UCC) visits
4. Updated COVID-19-related controls to account for pandemic-induced regional differences in 2021

To assess whether COVID-19 affected outcomes differentially between CPC+ and comparison regions in PY 4 and PY 5, we studied the direct effects of COVID-19 as well as overall changes in health care utilization resulting from a combination of direct and indirect effects. We measured direct effects by examining the prevalence of COVID-19 and related diagnoses in Medicare FFS claims and excess deaths due to the pandemic. Indirect effects refer to impacts caused by behavioral response to the COVID-19 pandemic, such as health care avoidance or hospitals suspending elective surgeries. Two regions with the same direct effect of COVID-19 could experience different behavioral responses. Therefore, to understand how COVID-19 could affect the CPC+ impact evaluation, it is important to examine the total effect of the pandemic, or the combination of direct and indirect effects, as captured by net changes in health care utilization. We do this by examining changes in key, acute care outcomes of the CPC+ impact evaluation with a difference-in-differences model, using 2019 (the year before the pandemic) as a "baseline" to check for larger-than-expected divergence in trends in 2020 and 2021. Finally, we assessed how adding COVID-19-related regional variables to the difference-in-differences model—the approach we adopted in AR4—accounted for regional differences due to COVID-19 in this year's analyses.

We implemented these analyses with three different populations: (1) the full Medicare fee-for-service (FFS) population living in CPC+ and comparison regions, (2) the intent-to-treat (ITT) CPC+ impact analysis sample, and (3) nonparticipating practices located in CPC+ regions and unselected practices located in comparison regions. We used the full Medicare FFS population to calculate excess death rates, which require a large, stable denominator. Since we are primarily interested in the effects of COVID-19 on the ITT analysis sample, we present results for this population. However, we were concerned about the ability to distinguish between the effects of COVID-19 and the effects of CPC+ for some health care utilization outcomes. Therefore, we also examined differences for unselected practices in CPC+ and comparison regions, which should not be directly affected by CPC+. We assumed that COVID-19 similarly affected selected and unselected practices in the same regions, and our findings presented later in the Appendix support this assumption.

5.D.2. Direct effects on COVID-19-related diagnoses and excess deaths

What's new this year for the direct effects?

As in AR4, we continued to look at COVID-19-related diagnoses and excess deaths in each region, now including 2021. We found:

1. COVID-19-related diagnosis rates were higher in 2021 than 2020, while rates of excess deaths were lower in 2021 than 2020.
2. Regional differences remained small across these outcomes in 2021.

A. COVID-19-related diagnoses

We estimated the rates of COVID-19-related diagnoses (see text box below for definition) among CPC+ and comparison beneficiaries as well as beneficiaries attributed to unselected practices in both types of regions (see Figure 5.D.i. in Supplement 1 for the regional distribution of CPC+ and comparison practices).[50] Using unselected practices allowed us to look at regional differences among beneficiaries who are not attributed to practices in the evaluation, leading to estimates that are not affected by CPC+.

Closer look: COVID-19-related diagnoses

COVID-19-related diagnoses include:

- COVID-19 diagnoses, identified by searching all primary and secondary diagnoses for the ICD-10 code B9729 (other coronavirus) before April 1, 2020, and U071 (2019 Novel Coronavirus) from April 1, 2020, onwards
- Respiratory conditions related to COVID-19, defined as claims with primary and secondary diagnoses for any of the following:
  - Viral pneumonia (J1289)
  - Bronchitis – acute (J208) or unspecified (J40)
  - Lower respiratory infection – specified (J988) or unspecified (J22)
  - Acute respiratory distress syndrome (J80)
  - Pneumonia because of COVID-19 (J12.82)

We included both COVID-19 diagnosis and COVID-19-related respiratory diagnoses to identify all cases that might have been caused by COVID-19, including cases misdiagnosed early in the pandemic.

Source: Bohl and Roozeboom-Baker (2020).

First, key findings from our analysis of COVID-19-related diagnoses are as follows:

- Rates of COVID-19-related diagnoses were slightly higher in 2021 compared to 2020 for all practice groups in both tracks (Table 5.D.1). Over the course of 2020, around 7 percent of CPC+ and comparison beneficiaries and 8 percent of non-CPC+ and non-comparison beneficiaries had a Medicare Part A or B claim with a COVID-19-related diagnosis in each track. In 2021, this number was 9 percent for CPC+ and comparison beneficiaries and 10 percent for non-CPC+ and non-comparison beneficiaries. From 2020 to 2021, the monthly rate of COVID-diagnoses increased from 1.0 to 1.2 percent of CPC+ and comparison beneficiaries.[51]

[50] Using claims data to identify COVID-19 cases could result in an undercount of true COVID-19-related diagnoses, given at-home testing and asymptomatic cases.

[51] Some beneficiaries had a COVID-19-related diagnosis during multiple months in 2020, which is why the cumulative percentage of CPC+ and comparison beneficiaries who had a COVID-19-related diagnosis between March and December 2020 are 7 percent and not 10 percent (1 percent multiplied by 10 months). Similarly, some beneficiaries had a COVID-19-related diagnosis during multiple months in 2021, so the cumulative percentage of CPC+ and comparison beneficiaries who had a COVID-19-related diagnosis between January and December 2021 are 9 percent and not 14.4 percent (1.2 percent multiplied by 12 months).

- CPC+ and comparison beneficiaries had similar average numbers of COVID-19-related diagnoses in 2020 and 2021 but there were differences in the timing of diagnoses, reflecting the geographic spread of COVID-19 across the country (Figure 5.D.ii in Supplement 2). For example, CPC+ practices had somewhat more beneficiaries diagnosed early in the pandemic (in April 2020) relative to practices in comparison regions and somewhat fewer beneficiaries diagnosed through the end of 2020. The differential timing of diagnoses in 2020 was somewhat more pronounced for Track 1 practices, likely driven by New Jersey, which was hit hard early in the pandemic and had a high proportion of Track 1 CPC+ practices. In 2021, the differences in diagnoses between CPC+ and comparison regions were small, particularly later in the year.
- Beneficiaries with a COVID-19-related diagnosis in the CPC+ and comparison groups were similar on characteristics such as age, race, original reason for Medicare entitlement, and chronic conditions (Table 5.D.i in Supplement 2).
- For both 2020 and 2021, the regional differences in the rate of COVID-19-related diagnoses were very similar in magnitude between selected and unselected practices (Table 5.D.1). In 2020, both selected and unselected practices in Track 1 had a 0.03 percentage point regional difference in the percentage of beneficiaries diagnosed each month. The differences were even smaller (less than 0.01 percentage points) in 2021. For Track 2, there were regional differences of less than 0.05 percentage points for selected and unselected practices each year; for unselected practices, the difference decreased from 2020 to 2021 (from 0.08 to 0.01 percentage points).

Second, key findings from our analysis of utilization and expenditures for COVID-19-related diagnoses are as follows:

- Rates of outpatient ED visits with COVID-19-related diagnoses were higher in 2021 than 2020 for all practice groups in both tracks, reflecting the higher rates of COVID-19-related diagnoses in 2021 (Table 5.D.1). For example, among CPC+ practices in both tracks, rates of outpatient ED visits with a COVID-19-related diagnosis increased from 10 per 1,000 beneficiaries in 2020 to 16 per 1,000 beneficiaries in 2021.
- By contrast, rates of and expenditures on acute hospitalizations with COVID-19-related diagnoses were either lower in 2021 than 2020 or stayed similar.
- For both 2020 and 2021, the regional differences in the utilization and expenditures for COVID-19-related diagnoses were small and similar in magnitude for both selected and unselected practices. For example, the regional differences in COVID-19-related acute hospitalizations were around half a percent of the rate of all-cause acute hospitalizations (Table 5.D.1).

Although we observed some larger regional differences within the SSP and non-SSP subgroups compared to those in the overall track, the differences in the rates of COVID-19-related diagnoses and utilization and expenditures for COVID-19-related diagnoses remained small (Tables 5.D.ii and 5.D.iii). For example, for both SSP and non-SSP practices, the regional differences in the rate of COVID-19-related diagnoses were less than 0.1 percentage points in 2020 and 2021. Similarly, the regional differences in acute hospitalizations for COVID-19-related diagnoses were less than 1 percent of the rate of all-cause acute hospitalizations in 2020 and 2021 for both SSP and non-SSP practices.

Overall, the direct effects of COVID-19 were small and similar in magnitude between selected and unselected practices in the CPC+ and comparison regions for both 2020 and 2021, in the overall sample and for SSP and non-SSP practices. The similarity of findings among selected and unselected practices suggests that, by examining outcomes among unselected practices, we are likely to accurately capture differential regional effects of the pandemic without including any effects of CPC+.

Table 5.D.1. COVID-19-related diagnoses and outpatient ED visits for COVID-19-related diagnoses were higher in 2021 than 2020, but COVID-19-related hospitalizations were somewhat lower in 2021 than 2020. Differences between practices in CPC+ and comparison regions were small both years.

	Track 1 – Unadjusted means				Track 1 – Differences (SE)		Track 2 – Unadjusted means				Track 1 – Differences (SE)	
	CPC+	Comparison	Non-CPC+	Non-comparison	CPC+ vs. comparison	Non-CPC+ vs. non-comparison	CPC+	Comparison	Non-CPC+	Non-comparison	CPC+ vs. comparison	Non-CPC+ vs. non-comparison
Beneficiary claims with COVID-19-related diagnosis (percentage of beneficiaries with a claim each month)												
2020 (Mar–Dec)	1.0%	1.0%	1.2%	1.2%	-0.03 p.p. (0.03 p.p.)	-0.03 p.p. (0.03 p.p.)	0.9%	1.0%	1.1%	1.2%	-0.03 p.p. (0.02 p.p.)	-0.08 p.p.*** (0.02 p.p.)
2021 (Jan–Dec)	1.2%	1.2%	1.4%	1.4%	0.00 p.p. (0.03 p.p.)	0.01 p.p. (0.02 p.p.)	1.2%	1.2%	1.3%	1.3%	0.05 p.p.* (0.03 p.p.)	0.01 p.p. (0.02 p.p.)
2020–2021[a]	1.1%	1.1%	1.3%	1.3%	-0.01 p.p. (0.02 p.p.)	-0.01 p.p. (0.02 p.p.)	1.1%	1.1%	1.2%	1.3%	0.01 p.p. (0.02 p.p.)	-0.03 p.p.* (0.02 p.p.)
Outpatient ED visits, including observation stays, with COVID-19-related diagnosis (per 1,000 beneficiaries per year)												
2020 (Mar–Dec)	10	12	12	14	-1.5*** (0.3)	-1.6*** (0.3)	10	11	12	14	-1.1*** (0.4)	-1.4*** (0.3)
2021 (Jan–Dec)	16	16	18	18	0.4 (0.4)	-0.1 (0.3)	16	15	19	18	0.4 (0.4)	0.4 (0.3)
2020–2021[a]	13	14	15	16	-0.5 (0.3)	-0.8*** (0.3)	13	13	16	16	-0.3 (0.4)	-0.4 (0.3)
Acute hospitalizations with COVID-19-related diagnosis (per 1,000 beneficiaries per year)												
2020 (Mar–Dec)	19	20	23	24	-1.2*** (0.4)	-0.7 (0.5)	18	20	22	24	-1.2** (0.5)	-1.4*** (0.4)
2021 (Jan–Dec)	19	19	22	21	0.1 (0.4)	0.8** (0.3)	19	18	21	21	0.2 (0.4)	0.8** (0.3)
2020–2021[a]	19	19	22	22	-0.4 (0.3)	0.1 (0.3)	19	19	22	22	-0.4 (0.4)	-0.2 (0.3)
Medicare inpatient expenditures for COVID-19-related diagnosis (per beneficiary per month)												
2020 (Mar–Dec)	$33	$37	$41	$45	-$3.9*** ($0.9)	-$3.5*** ($1.1)	$32	$36	$40	$45	-$3.7*** ($1.1)	-$4.8*** ($1.0)
2021 (Jan–Dec)	$34	$35	$40	$40	-$1.7** ($0.7)	-$0.2 ($0.7)	$33	$34	$39	$39	-$0.8 ($0.7)	$0.1 ($0.7)
2020–2021[a]	$33	$36	$41	$42	-$2.6*** ($0.7)	-$1.7** ($0.7)	$33	$35	$40	$42	-$2.1*** ($0.8)	-$2.2*** ($0.7)

Table 5.D.1. ***(continued)***

	Track 1 – Unadjusted means				Track 1 – Differences (SE)		Track 2 – Unadjusted means				Track 1 – Differences (SE)	
	CPC+	Comparison	Non-CPC+	Non-comparison	CPC+ vs. comparison	Non-CPC+ vs. non-comparison	CPC+	Comparison	Non-CPC+	Non-comparison	CPC+ vs. comparison	Non-CPC+ vs. non-comparison
Unweighted sample sizes												
Number of practices	1,373	5,242	8,335	20,654			1,515	3,783	7,274	20,113		
Average number of beneficiaries per month	1,027,990	3,589,897	2,232,413	6,694,166			1,263,737	3,032,104	1,826,954	6,518,590		

Source: Mathematica's analysis of Medicare claims data from March 2020 through December 2021.

Note: COVID-19-related diagnoses include COVID-19 diagnoses and respiratory conditions related to COVID-19 including viral pneumonia, acute bronchitis, lower respiratory infection, acute respiratory distress syndrome, and pneumonia due to COVID-19. See Bohl and Roozeboom-Baker (2020) for details. Differences in the table are from time-series models run at the practice-month-year level that did not adjust for beneficiary or practice characteristics. For CPC+ practices, observations were weighted by the number of Medicare FFS beneficiaries assigned to the practice during the month and year. For comparison practices, the weight is a product of the number of assigned beneficiaries and the matching weight. For non-CPC+ and non-comparison practices, we used a concentration weight constructed at the state-HRR level such that non-CPC+ practices had the same level of representation (in terms of beneficiary months) as CPC+ practices in the same state and HRR and SSP group, and non-comparison practices had the same level of representation as comparison practices in the same state and HRR and SSP group. Among non-CPC+ and non-comparison practices, we winsorized the weights at the 99th percentile.

[a] The 2020–2021 estimates used data from March 2020 to December 2021.

*/**/*** Significantly different from zero at the 0.10/0.05/0.01 levels, two-tailed test.

CPC+ = Comprehensive Primary Care Plus; ED = emergency department; HRR = hospital referral region; p.p. = percentage points; SE = standard error; SSP = Medicare Shared Savings Program.

B. Excess deaths

We estimated excess deaths as the difference between observed deaths in each month from March 2020 through December 2021 and predicted deaths during those same months if COVID-19 had not occurred. We used enrollment data for all Medicare FFS beneficiaries living in CPC+ and comparison regions to identify historical trends in deaths between 2016 and 2019 and then projected these trends out to months in 2020 and 2021 to estimate predicted deaths if COVID-19 had not occurred. This approach is consistent with methods to calculate excess deaths used in the COVID-19 literature (Polyakova et al. 2020). For calculating excess deaths, we weighted the population of all Medicare FFS beneficiaries in CPC+ regions to look similar to CPC+ beneficiaries in terms of age, race, sex, and location, with separate weighting by track and by SSP status. Taking the same approach, we also weighted all Medicare FFS beneficiaries in comparison regions to look similar to comparison beneficiaries. A detailed description of our methods for estimating excess deaths is available in Supplement 3.

Key findings from our analysis of excess deaths are as follows:

- CPC+ and comparison regions had similar average monthly excess deaths during March 2020 through December 2021 (Table 5.D.2). When weighted to represent Track 1 practices, deaths in CPC+ and comparison regions increased by 17 and 18 percent, respectively, compared to predicted deaths; when weighted to represent Track 2 practices, these increases were 16 and 17 percent, respectively, compared to predicted deaths.
- The average excess deaths during March to December 2020 (an increase of approximately 20 percent compared to predicted deaths) were somewhat higher than in 2021 (an increase of approximately 12 percent compared to predicted deaths) when weighted to represent either track (Table 5.D.2), likely driven by the availability of COVID-19 vaccines, more effective treatment strategies, and less deadly virus variants in 2021.
- Across 2020 and 2021, CPC+ regions had 0.1 fewer deaths per 10,000 beneficiaries per month than comparison regions for both tracks and these differences were not statistically distinguishable from zero (Table 5.D.2).
- Although there were no regional differences in average excess deaths for each pandemic year, there were some differences in individual months. Relative to comparison regions, CPC+ regions had greater increases in excess deaths in some months (for example, April 2020) and smaller increases in other months (for example, January 2021), reflecting geographic differences in the timing of the surges in the COVID-19 pandemic. These monthly differences were typically small (less than 3 percent of the predicted deaths in comparison regions) (data not shown).
- Results were similar when practices were weighted to represent the SSP or non-SSP subgroup within each track (Table 5.D.iv in Supplement 3).

Differences in timing of excess deaths are notable because they could have initiated differential responses to the pandemic that differentially affect the CPC+ evaluation outcomes, which we explore in later sections on indirect effects of the pandemic. This difference in timing of excess deaths was particularly pronounced in 2020 among regions weighted to represent Track 1 practices (see Figure 5.D.iii. in Supplement 3). Differences in timing appear to be driven—at least in part—by CPC+ practices in New Jersey (data not shown). In later sections, we explore using regional excess deaths as a control variable to account for these regional differences (see Supplement 4 for a description of methods we used to develop an excess deaths regional control variable).

Table 5.D.2. Excess deaths in 2020 and 2021 were similar among Medicare FFS beneficiaries in CPC+ and comparison regions[a]

Medicare FFS beneficiaries weighted to represent	Average monthly excess deaths in March 2020–December 2021, in deaths per 10,000 beneficiaries per month (Percentage change from historical trends[a])			
	CPC+ regions[b]	Comparison regions[b]	Difference	90% confidence interval
Track 1				
2020 (March–December)	7.0 (20%)	7.1 (21%)	0.0	(-0.9, 0.8)
2021 (January–December)	4.3 (12%)	4.6 (13%)	-0.3	(-0.9, 0.4)
2020–2021c	6.1 (17%)	6.2 (18%)	-0.1	(-0.7, 0.4)
Track 2				
2020 (March–December)	6.6 (19%)	6.7 (20%)	-0.1	(-1.1, 0.8)
2021 (January–December)	4.4 (12%)	4.4 (13%)	-0.1	(-0.7, 0.6)
2020–2021c	5.8 (16%)	5.9 (17%)	-0.1	(-0.8, 0.6)

Source: Mathematica's analysis of Medicare enrollment data from January 2016 through December 2021.

Note: Excess deaths are the difference between observed deaths in March 2020 through December 2021 and predicted deaths if COVID-19 had not occurred. Predicted deaths are based on models that are regression-adjusted for the distribution of age, race, and sex in the region. The models use data from 2016 through 2019 and project trends out through 2021 to predict deaths if the COVID-19 pandemic had not occurred. For calculating excess deaths, we used observations at the state and HRR, month, year, age group, race, and sex levels. Each observation was weighted based on (1) the share of the 2019 ITT sample of CPC+ or comparison beneficiaries in that state-HRR, by track; (2) the share of the 2019 ITT sample of CPC+ or comparison beneficiaries in that age-race-sex cell, by track; and (3) the matching weights of comparison group practices in a state-HRR, by track. For a detailed description of methods, see Supplement 3.

[a] To calculate these percentages, we divided the excess deaths in the region by the predicted deaths if COVID-19 had not occurred. Predicted deaths are based on regression models using data from 2016 through 2019.

[b] Regions defined as the combination of state and HRR.

[c] The 2020–2021 estimates used data from March 2020 to December 2021.

CPC+ = Comprehensive Primary Care Plus; FFS = fee for service; HRR = hospital referral region; ITT = intent to treat.

5.D.3. Total effects of COVID-19 on health care utilization and Medicare expenditures

What's new this year for the total effects?

As in AR4, we continued to look at changes in key Medicare outcomes among unselected practices in CPC+ and comparison regions, now including 2021 data and additional outcomes, and stratifying by SSP status. We found:

1. Utilization and expenditures began to return to their pre-pandemic levels in 2021 after their large decline in 2020, except for UCC visits which grew over the course of the pandemic.
2. Regional differences remained small across our outcomes and practice groups in 2021 (less than 3-percent differences), with a few exceptions.

We examined total effects of COVID-19, that is the combination of direct and indirect effects, by measuring changes in key outcomes of the CPC+ impact evaluation between 2019 and 2021. We studied Medicare Part A and B expenditures and expenditures on acute hospitalizations and outpatient ED visits. In addition, we studied acute care utilization, specifically acute hospitalizations, outpatient ED visits, and UCC visits. We examined these outcomes because they are the key outcomes of interest in the CPC+ impact evaluation and have high potential to be influenced by COVID-19. To approximate regional

changes due to the pandemic without confounding those with the effects of CPC+, we examined net changes in health care utilization and expenditures among unselected practices in CPC+ and comparison regions (that is, non-CPC+ and non-comparison practices).

Key findings on the total effects among unselected practices were as follows:

- Over the course of the pandemic (2020 and 2021 combined), unadjusted means for nearly all outcomes were lower than in 2019, with larger declines in 2020 than 2021 (Tables 5.D.3 and 5.D.4). In the case of Medicare expenditures, there was a 30 percent dip in April 2020 before expenditures rebounded (Figure 5.D.vii in Supplement 7). Unlike the other outcomes, Medicare expenditures in 2021 surpassed 2019 levels by 2 percent for non-CPC+ practices in both tracks and by 4 and 3 percent for non-comparison practices in Tracks 1 and 2, respectively.
- Declines in utilization and expenditures in 2020 and 2021 (relative to 2019) were larger among non-CPC+ practices compared to non-comparison practices, but regional differences were small for most outcomes (less than 2 percent) (Tables 5.D.3 and 5.D.4). Also, for most outcomes, regional differences were roughly similar in 2020 and 2021, with some exceptions. In both tracks, regional differences in the rates of acute hospitalizations and expenditures for acute hospitalizations were higher in 2020 than in 2021. Similarly, for Medicare total expenditures in Track 2, regional differences were larger in 2021 than in 2020.
- UCC visits differed from the general trends for other outcomes, with increasing rates during the pandemic for both CPC+ and comparison practices and larger increases among non-CPC+ practices, relative to non-comparison practices (Tables 5.D.3 and 5.D.4; Figure 5.D.viii in Supplement 7). Over the course of the pandemic, rates of UCC visits grew substantially from their 2019 levels for non-CPC+ practices (30 percent in Track 1 and 24 percent in Track 2). UCC visits also grew for non-comparison practices, but to a lesser extent (by 17 and 13 percent in Tracks 1 and 2). The increases in UCC visits were especially large in 2021 for both tracks, for example, increasing by 52 percent for non-CPC+ practices (Track 1) and 36 percent for non-comparison practices (Track 1). The observed growth in UCC visits is consistent with the documented surge in the number of UCCs nationwide in recent years (57 percent growth nationally between 2013 and 2019), combined with UCCs becoming a common site to obtain COVID-19-related care, such as testing and vaccinations (Urgent Care Association 2019, n.d.; Yousman et al. 2021).
- When we stratified non-CPC+ and non-comparison practices by SSP status, the regional differences between practices with the same SSP status were larger compared to the overall track differences, but were still typically less than 3 percent (Tables 5.D.3 and 5.D.4). Two exceptions to this pattern in both tracks were expenditures on outpatient ED visits in SSP practices (differences up to 6 percent in Track 2) and rates of UCC visits in both SSP and non-SSP practices (differences up to 16 percent among Track 1 SSP practices).
- The observed unadjusted regional differences among unselected practices in changes from 2019 to 2020 and from 2019 to 2021 were not substantially greater than the normal year-to-year and two-year–to–two-year variation, respectively, that occurred prior to the pandemic (Figure 5.D.1 below and Figure 5.D.ix in Supplement 7). The one exception is UCC visits, where there was substantial growth during the pandemic relative to historical variation (Panel E of Figure 5.D.ix in Supplement 7). We expect to see variation when comparing estimates over time for reasons unrelated to COVID-19, including random variation and other secular trends. We expect that variation to be greater when comparing across two years, than when comparing across just one year.

Table 5.D.3. Non-CPC+ practices had greater decreases in health care utilization and expenditures than non-comparison practices between 2019 and 2021: Unadjusted results (Track 1)

	Track 1 All				Track 1 SSP				Track 1 non-SSP			
	Non-CPC+ mean	Non-comparison mean	Non-CPC+ vs. non-comparison differences relative to 2019 (SE)	Percentage difference relative to 2019 non-comparison mean[a]	Non-CPC+ mean	Non-comparison mean	Non-CPC+ vs. non-comparison differences relative to 2019 (SE)	Percentage difference relative to 2019 non-comparison mean[a]	Non-CPC+ mean	Non-comparison mean	Non-CPC+ vs. non-comparison differences relative to 2019 (SE)	Percentage difference relative to 2019 non-comparison mean[a]
Medicare Part A and B expenditures without enhanced payments for CPC+ and SSP												
2019 (Jan–Dec)	$1,045	$1,059	NA	NA	$1,066	$1,080	NA	NA	$1,023	$1,034	NA	NA
2020 (Jan–Dec)	$994	$1,019	-$10.6*** ($3.8)	-1.0%	$1,015	$1,032	-$4.0 ($6.2)	-0.4%	$973	$1,002	-$18.5*** ($4.6)	-1.8%
2021 (Jan–Dec)	$1,071	$1,097	-$11.4** ($4.5)	-1.1%	$1,100	$1,110	$3.3 ($7.2)	0.3%	$1,040	$1,080	-$28.7*** ($5.4)	-2.8%
2020–2021 (combined COVID-19 years)	$1,032	$1,058	-$11.0*** ($3.7)	-1.0%	$1,056	$1,070	-$0.4 ($6.0)	-0.04%	$1,006	$1,040	-$23.5*** ($4.3)	-2.3%
Acute hospitalizations (short-stay acute care and critical access hospitals)												
2019 (Jan–Dec)	312	303	NA	NA	315	303	NA	NA	310	302	NA	NA
2020 (Jan–Dec)	264	259	-5.1*** (1.5)	-1.7%	267	260	-5.1** (2.5)	-1.7%	260	259	-6.1*** (1.8)	-2.0%
2021 (Jan–Dec)	262	255	-2.6 (1.7)	-0.8%	267	255	0.0 (2.8)	0.00%	256	255	-6.2*** (2.0)	-2.1%
2020–2021 (combined COVID-19 years)	263	257	-3.8*** (1.4)	-1.3%	267	257	-2.6 (2.4)	-0.9%	258	257	-6.1*** (1.7)	-2.0%
Expenditures on acute hospitalizations (short-stay acute care and critical access hospitals) (per beneficiary per month)												
2019 (Jan–Dec)	$327	$332	NA	NA	$333	$337	NA	NA	$320	$325	NA	NA
2020 (Jan–Dec)	$307	$318	-$5.3** ($2.2)	-1.6%	$314	$322	-$4.0 ($3.3)	-1.2%	$301	$313	-$7.5*** ($2.7)	-2.3%
2021 (Jan–Dec)	$319	$327	-$2.5 ($2.3)	-0.8%	$327	$328	$3.3 ($3.6)	1.0%	$309	$324	-$9.8*** ($2.8)	-3.0%
2020–2021 (combined COVID-19 years)	$313	$322	-$3.9** ($1.9)	-1.2%	$321	$325	-$0.4 ($3.1)	-0.1%	$305	$318	-$8.6*** ($2.3)	-2.7%
Outpatient ED visits, including observation stays												
2019 (Jan–Dec)	524	531	NA	NA	509	509	NA	NA	543	550	NA	NA
2020 (Jan–Dec)	404	416	-4.6* (2.4)	-0.9%	388	398	-10.0** (3.9)	-2.0%	424	433	-1.2 (3.0)	-0.2%
2021 (Jan–Dec)	431	442	-4.3 (2.9)	-0.8%	418	424	-6.7 (4.3)	-1.3%	448	460	-4.6 (3.7)	-0.8%
2020–2021 (combined COVID-19 years)	417	429	-4.5* (2.4)	-0.8%	402	411	-8.4** (3.7)	-1.7%	436	446	-2.8 (3.1)	-0.5%

Table 5.D.3. ***(continued)***

	Track 1 All				Track 1 SSP				Track 1 non-SSP			
	Non-CPC+ mean	Non-comparison mean	Non-CPC+ vs. non-comparison differences relative to 2019 (SE)	Percentage difference relative to 2019 non-comparison mean[a]	Non-CPC+ mean	Non-comparison mean	Non-CPC+ vs. non-comparison differences relative to 2019 (SE)	Percentage difference relative to 2019 non-comparison mean[a]	Non-CPC+ mean	Non-comparison mean	Non-CPC+ vs. non-comparison differences relative to 2019 (SE)	Percentage difference relative to 2019 non-comparison mean[a]
Expenditures on outpatient ED visits, including observation stays (per beneficiary per month)												
2019 (Jan–Dec)	$31	$33	NA	NA	$30	$31	NA	NA	$32	$34	NA	NA
2020 (Jan–Dec)	$26	$28	-$0.3 ($0.2)	-0.9%	$24	$26	-$0.7** ($0.3)	-2.4%	$28	$29	$0.1 ($0.3)	0.3%
2021 (Jan–Dec)	$30	$32	-$0.5* ($0.3)	-1.5%	$28	$31	-$1.4*** ($0.4)	-4.5%	$32	$33	$0.3 ($0.4)	0.9%
2020–2021 (combined COVID-19 years)	$28	$30	-$0.4* ($0.2)	-1.2%	$26	$28	-$1.1*** ($0.3)	-3.4%	$30	$31	$0.2 ($0.3)	0.6%
Urgent care center visits (per 1,000 beneficiaries per year)												
2019 (Jan–Dec)	124	129	NA	NA	136	140	NA	NA	112	119	NA	NA
2020 (Jan–Dec)	135	128	12.7*** (1.7)	9.8%	151	139	16.1*** (3.0)	11.5%	118	116	9.4*** (1.8)	7.9%
2021 (Jan–Dec)	188	175	18.3*** (2.9)	14.1%	213	195	22.1*** (5.0)	15.8%	160	154	13.6*** (2.9)	11.4%
2020–2021 (combined COVID-19 years)	161	151	15.4*** (2.1)	11.9%	181	166	19.0*** (3.7)	13.6%	139	135	11.5*** (2.2)	9.7%
Unweighted sample sizes												
Number of practices	8,337	20,654			2,488	5,151			5,849	15,503		
Average number of beneficiaries per month	2,284,100	6,798,349			809,545	2,167,584			1,474,555	4,630,765		

Source: Mathematica's analysis of Medicare claims data from January 2019 through December 2021.

Note: Differences in the table are from time-series models run at the practice-month-year level that did not adjust for beneficiary or practice characteristics. For these practices, we used a concentration weight constructed at the state-HRR level such that non-CPC+ practices had the same level of representation (in terms of beneficiary months) as CPC+ practices in the same state and HRR and SSP group, and non-comparison practices had the same level of representation as comparison practices in the same state and HRR and SSP group. We winsorized the weights at the 99th percentile. Standard errors are clustered at the practice level. For a detailed description of methods, see Supplement 5.

*/**/*** Significantly different from zero at the 0.10/0.05/0.01 levels, two-tailed test.

[a] To calculate these percentages, we divided the non-CPC+ vs. non-comparison differences by the unadjusted 2019 non-comparison mean for the outcome.

CPC+ = Comprehensive Primary Care Plus; ED = emergency department; FFS = fee-for-service; HRR = hospital referral region; NA = not applicable; SE = standard error; SSP = Medicare Shared Savings Program.

Table 5.D.4. Non-CPC+ practices had greater decreases in health care utilization and expenditures than non-comparison practices between 2019 and 2021: Unadjusted results (Track 2)

	Track 2 All				Track 2 SSP				Track 2 non-SSP			
	Non-CPC+ mean	Non-comparison mean	Non-CPC+ vs. non-comparison differences relative to 2019 (SE)	Percentage difference relative to 2019 non-comparison mean[a]	Non-CPC+ mean	Non-comparison mean	Non-CPC+ vs. non-comparison differences relative to 2019 (SE)	Percentage difference relative to 2019 non-comparison mean[a]	Non-CPC+ mean	Non-comparison mean	Non-CPC+ vs. non-comparison differences relative to 2019 (SE)	Percentage difference relative to 2019 non-comparison mean[a]
Medicare Part A and B expenditures without enhanced payments for CPC+ and SSP												
2019 (Jan–Dec)	$1,036	$1,057	NA	NA	$1,063	$1,087	NA	NA	$1,015	$1,026	NA	NA
2020 (Jan–Dec)	$986	$1,015	-$8.2** ($4.0)	-0.8%	$1,006	$1,039	-$8.9 ($8.7)	-0.8%	$970	$991	-$9.0** ($4.5)	-0.9%
2021 (Jan–Dec)	$1,058	$1,091	-$12.2*** ($4.7)	-1.2%	$1,084	$1,113	-$5.0 ($9.6)	-0.5%	$1,036	$1,068	-$19.7*** ($5.3)	-1.9%
2020–2021 (combined COVID-19 years)	$1,021	$1,053	-$10.2*** ($3.9)	-1.0%	$1,045	$1,075	-$7.0 ($8.5)	-0.6%	$1,002	$1,028	-$14.3*** ($4.3)	-1.4%
Acute hospitalizations (short-stay acute care and critical access hospitals)												
2019 (Jan–Dec)	312	306	NA	NA	327	309	NA	NA	301	302	NA	NA
2020 (Jan–Dec)	265	261	-2.6* (1.5)	-0.9%	279	265	-3.7 (3.3)	-1.2%	255	258	-2.4 (1.7)	-0.8%
2021 (Jan–Dec)	262	256	-0.2 (1.7)	-0.1%	279	258	3.8 (3.5)	1.2%	250	254	-3.4* (1.9)	-1.1%
2020–2021 (combined COVID-19 years)	264	259	-1.4 (1.5)	-0.5%	279	261	0.0 (3.2)	-0.01%	252	256	-2.9* (1.6)	-1.0%
Expenditures on acute hospitalizations (short-stay acute care and critical access hospitals) (per beneficiary per month)												
2019 (Jan–Dec)	$327	$334	NA	NA	$334	$342	NA	NA	$321	$324	NA	NA
2020 (Jan–Dec)	$308	$319	-$3.3 ($2.2)	-1.0%	$316	$327	-$2.7 ($4.1)	-0.8%	$302	$309	-$4.0 ($2.7)	-1.2%
2021 (Jan–Dec)	$320	$326	$1.6 ($2.4)	0.5%	$332	$332	$8.4** ($4.3)	2.5%	$311	$319	-$4.7 ($2.9)	-1.4%
2020–2021 (combined COVID-19 years)	$314	$322	-$0.9 ($2.0)	-0.3%	$324	$329	$2.8 ($3.7)	0.8%	$307	$314	-$4.3* ($2.4)	-1.3%
Outpatient ED visits, including observation stays												
2019 (Jan–Dec)	547	534	NA	NA	522	518	NA	NA	567	548	NA	NA
2020 (Jan–Dec)	421	419	-11.0*** (2.3)	-2.1%	397	408	-14.6*** (4.2)	-2.8%	441	431	-8.6*** (2.9)	-1.6%
2021 (Jan–Dec)	446	447	-13.8*** (2.9)	-2.6%	424	434	-14.3*** (4.8)	-2.8%	464	458	-12.6*** (3.8)	-2.3%
2020–2021 (combined COVID-19 years)	434	433	-12.3*** (2.4)	-2.3%	410	420	-14.5*** (4.0)	-2.8%	453	444	-10.6*** (3.1)	-1.9%

Table 5.D.4. ***(continued)***

	Track 2 All				Track 2 SSP				Track 2 non-SSP			
	Non-CPC+ mean	Non-comparison mean	Non-CPC+ vs. non-comparison differences relative to 2019 (SE)	Percentage difference relative to 2019 non-comparison mean[a]	Non-CPC+ mean	Non-comparison mean	Non-CPC+ vs. non-comparison differences relative to 2019 (SE)	Percentage difference relative to 2019 non-comparison mean[a]	Non-CPC+ mean	Non-comparison mean	Non-CPC+ vs. non-comparison differences relative to 2019 (SE)	Percentage difference relative to 2019 non-comparison mean[a]
Expenditures on outpatient ED visits, including observation stays (per beneficiary per month)												
2019 (Jan–Dec)	$32	$33	NA	NA	$31	$33	NA	NA	$33	$33	NA	NA
2020 (Jan–Dec)	$27	$28	-$0.7*** ($0.2)	-2.2%	$25	$28	-$1.4*** ($0.4)	-4.2%	$28	$28	-$0.3 ($0.3)	-0.9%
2021 (Jan–Dec)	$31	$32	-$0.8*** ($0.3)	-2.4%	$29	$32	-$2.0*** ($0.4)	-6.2%	$33	$32	$0.0 ($0.4)	-0.1%
2020–2021 (combined COVID-19 years)	$29	$30	-$0.8*** ($0.2)	-2.3%	$27	$30	-$1.7*** ($0.3)	-5.2%	$30	$30	-$0.2 ($0.3)	-0.5%
Urgent care center visits (per 1,000 beneficiaries per year)												
2019 (Jan–Dec)	123	127	NA	NA	130	135	NA	NA	118	119	NA	NA
2020 (Jan–Dec)	128	123	8.6*** (1.6)	6.8%	135	132	8.6*** (3.1)	6.3%	123	115	8.9*** (2.0)	7.5%
2021 (Jan–Dec)	177	165	15.3*** (2.7)	12.1%	187	178	14.3*** (4.6)	10.6%	170	152	17.7*** (3.2)	14.9%
2020–2021 (combined COVID-19 years)	152	144	11.9*** (2.0)	9.4%	161	154	11.4*** (3.5)	8.4%	146	133	13.2*** (2.4)	11.1%
Unweighted sample sizes												
Number of practices	7,276	20,113			2,423	5,010			4,853	15,103		
Average number of beneficiaries per month	1,871,011	6,619,682			756,018	2,094,775			1,114,993	4,524,907		

Source: Mathematica's analysis of Medicare claims data from January 2019 through December 2021.

Note: Differences in the table are from time-series models run at the practice-month-year level that did not adjust for beneficiary or practice characteristics. For these practices, we used a concentration weight constructed at the state-HRR level such that non-CPC+ practices had the same level of representation (in terms of beneficiary months) as CPC+ practices in the same state and HRR and SSP group, and non-comparison practices had the same level of representation as comparison practices in the same state and HRR and SSP group. We winsorized the weights at the 99th percentile. Standard errors are clustered at the practice level. For a detailed description of methods, see Supplement 5.

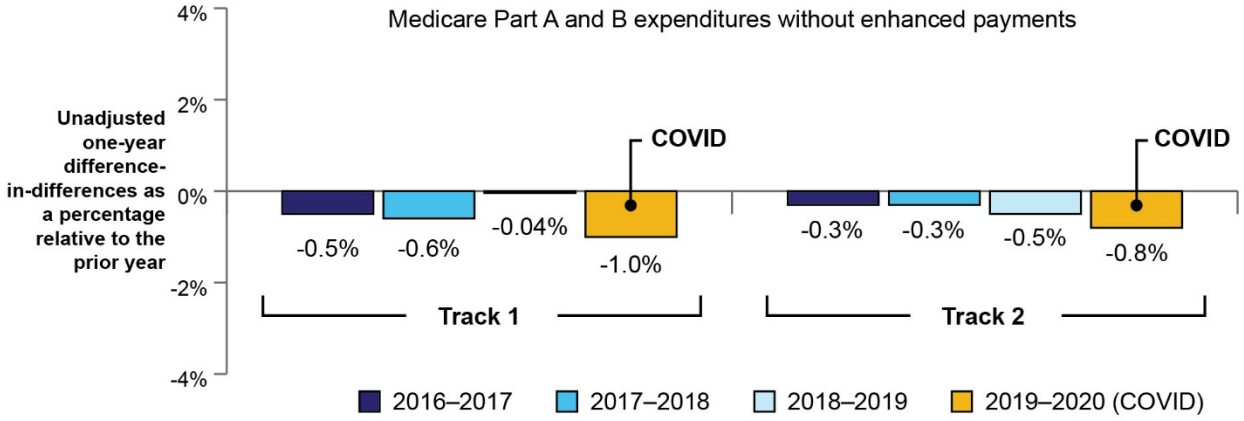
*/**/*** Significantly different from zero at the 0.10/0.05/0.01 levels, two-tailed test.

[a] To calculate these percentages, we divided the non-CPC+ vs. non-comparison differences by the unadjusted 2019 non-comparison mean for the outcome.

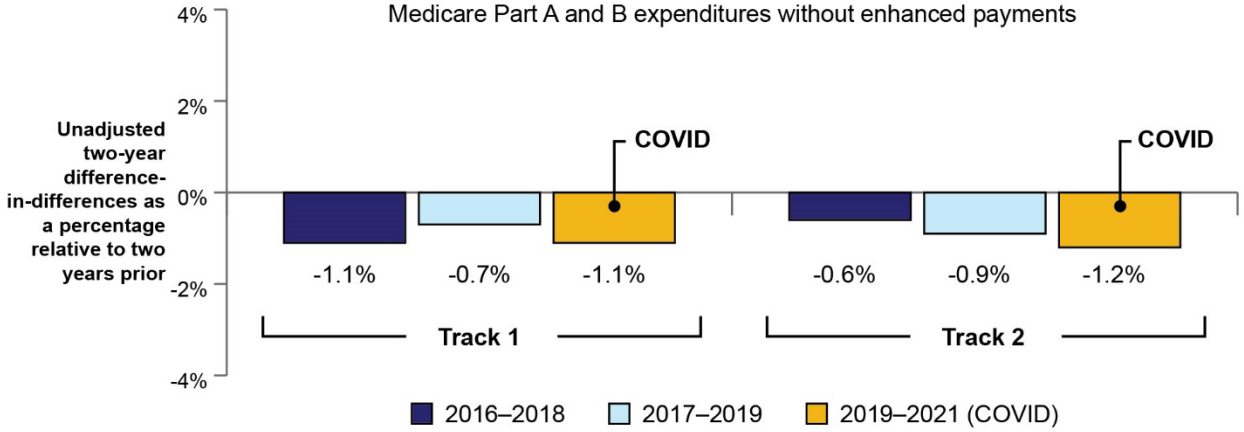
CPC+ = Comprehensive Primary Care Plus; ED = emergency department; FFS = fee-for-service; HRR = hospital referral region; NA = not applicable; SE = standard error; SSP = Medicare Shared Savings Program.

Figure 5.D.1. Regional differences in the change of Medicare expenditures among unselected practices from 2019 to 2020 and 2019 to 2021 were not substantially greater than historical changes

Panel A. Year-to-year variation among non-CPC+ practices relative to non-comparison practices



Panel B. Two-year–to–two-year variation among non-CPC+ practices relative to non-comparison practices



Source: Mathematica's analysis of Medicare claims data from January 2016 through December 2021.

Note: The figure shows the unadjusted difference in the year-to-year (Panel A) and two-year–to–two-year (Panel B) variation (in percentage terms) for Medicare expenditures between non-CPC+ and non-comparison practices from 2016 through 2021. In each panel, we compared the variation during the pandemic period (that is, bars with blue shading and having "COVID" in their legend labels) with pre-pandemic, historical variation (that is, bars with red shading). For year-to-year variation, we did not include the 2020–2021 variation in the post-pandemic period because none of our regression analyses are based on this year-to-year variation. For two-year–to–two-year variation, we omitted the 2018–2020 variation in the pre-pandemic period because it contains a year during the pandemic period and is, therefore, less helpful to serve as a historical benchmark for the 2019 to 2021 change. For these practices, we used a concentration weight constructed at the state-HRR level such that non-CPC+ practices had the same level of representation (in terms of beneficiary months) as CPC+ practices in the same state and HRR and SSP group, and non-comparison practices had the same level of representation as comparison practices in the same state and HRR and SSP group. We winsorized the weights at the 99th percentile.

CPC+ = Comprehensive Primary Care Plus; ED = emergency department; HRR = hospital referral region; SSP = Medicare Shared Savings Program.

Although we focused on regional changes in health care utilization and expenditures to capture total effects of COVID-19, we wanted to rule out differential changes in sample composition among CPC+ versus comparison practices due to the pandemic. By changing patterns of health care utilization, the pandemic could potentially (1) affect the mix of patients attributed to CPC+ versus comparison practices, and (2) lead to differential rates of practice closure in CPC+ versus the comparison group due to shortfalls in practice revenue from health care avoidance during the pandemic. We did not find any evidence that the sample of practices changed meaningfully in 2020 or 2021 in either the CPC+ or the comparison group, or that there were differential changes in beneficiary sample composition for the CPC+ group relative to the comparison group (see Supplement 8, Sections A and B, for results of this analysis).

5.D.4. Using COVID-19-related control variables to account for regional differences due to COVID-19

What's new this year in accounting for COVID-19?

As in AR4, we continued to compare regional differences with and without including COVID-19 control variables. For AR5, we added in 2021 data, updated COVID-19-related controls for 2021, included additional outcomes, and stratified analyses by SSP status within each track. We found:

1. COVID-19-related controls reduced or had little effect on regional differences in 2020, but had less of a consistent effect in 2021.
2. After including the COVID-19-related controls, regional differences remained small across our outcomes and practice groups in 2021 (less than 3-percent differences), with a few exceptions.

Although the magnitude of the differential decline in health care utilization between 2019 and 2021 for CPC+ versus comparison regions was typically small (less than 2 percent), it could potentially bias the impact estimate for CPC+ in PY 4 and PY 5. This is because the impact analyses rely on detecting differential changes in outcomes between baseline and each intervention year for CPC+ practices versus the matched comparison practices. Therefore, without additional mechanisms to account for the effects of COVID-19 across regions, the impact estimate for PY 4 and PY 5 would potentially include pandemic-induced regional changes.

Following the same approach as in AR4, we used regional COVID-19 controls to examine whether and to what extent including these additional control variables reduces the estimated regional differences from COVID-19 during 2020 and 2021. The control variables measure underlying health status and vulnerability to COVID-19, as well as resilience and mitigation efforts that we do not capture with other controls included in the CPC+ impact evaluation.

A. COVID-19-related control variables for AR5

We used the same type of COVID-19 controls as in AR4 with an additional year of data for 2021. These regional COVID-19 control variables are presented in Table 5.D.5, with more details in Supplement 5. All COVID-19 regional controls are merged to a practice based on a practice's geographic location. Each control variable is based on either the calendar year or the time period that corresponded to a particular

wave of the pandemic. For the latter, we defined three waves of the pandemic in 2020, as in AR4. With the additional year of data for AR5, we extended the definition of pandemic waves to include four additional waves in 2021. A visualization of the wave definitions and how they correspond with national COVID-19 cases is presented in Supplement 5, Figure 5.D.v.

Table 5.D.5. Proposed COVID-19-related regional controls for the Fifth Annual Report impacts model

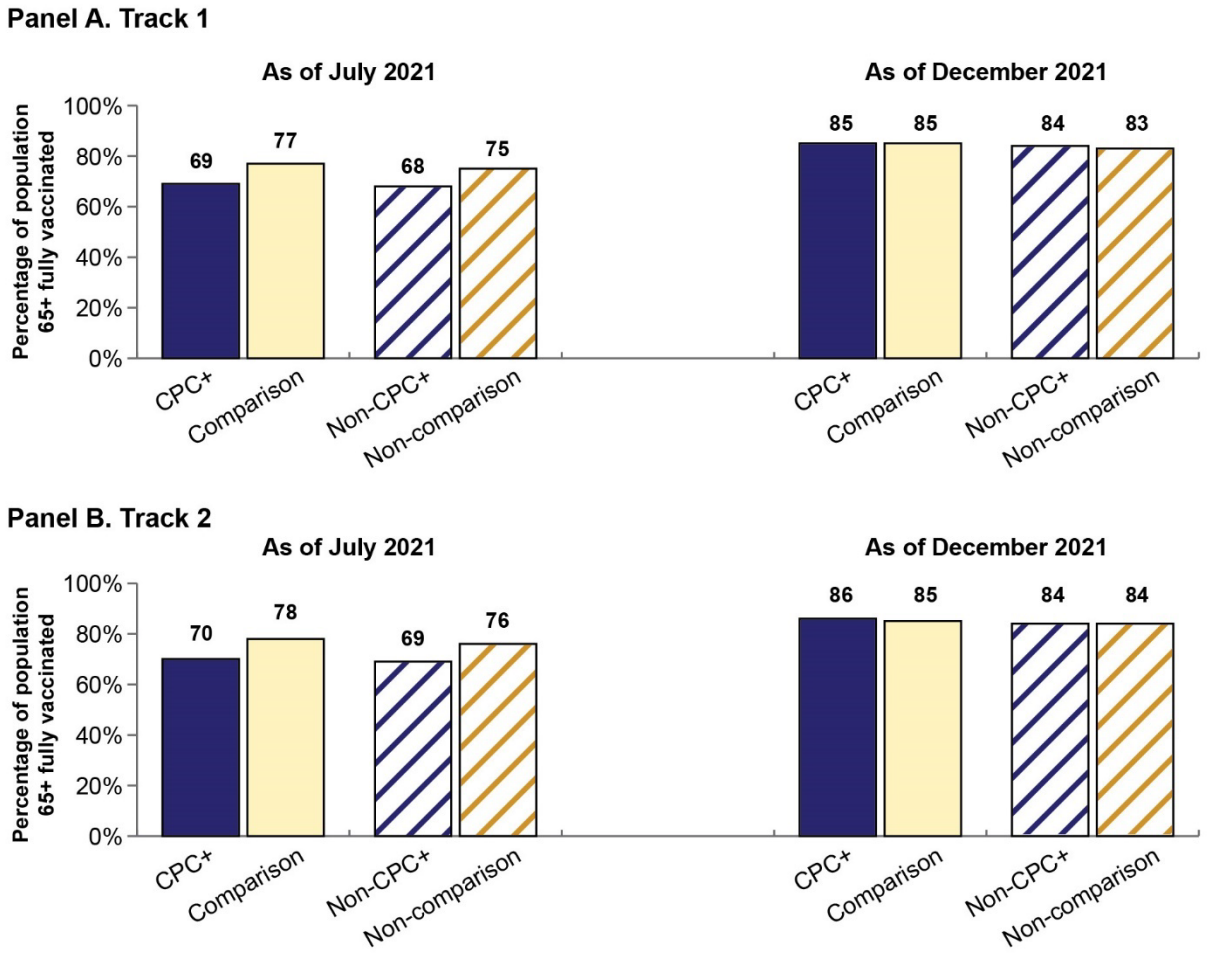
Control	Description	Rationale for Inclusion
Excess deaths	• Excess deaths refer to the number of all-cause deaths above-and-beyond what we would have predicted given historical trends. • Excess deaths in 2020 and 2021 indicate the severity of COVID-19 in the region during each year. • Using Bayesian methods to produce accurate estimates of regional excess deaths and predictive modeling methods consistent with the recent COVID-19 literature (Polyakova et al. 2020), we created a measure of excess deaths for 332 state-HRRs containing CPC+ or comparison practices. See Supplement 4 for more details. **Year(s) available:** 2020 and 2021 **Frequency:** Monthly **Geographic level for which the variable is defined:** State- HRR **Specific variable definitions included in model testing:** 1. *Average excess deaths in a wave:* We averaged excess monthly deaths during each "wave" of the pandemic for the state-HRR and interacted it with the contemporaneous calendar year (e.g., indicator for 2020 or 2021). 2. *Peak excess deaths per year:* We took the highest excess monthly death value in a calendar year for a state-HRR and interacted it with the contemporaneous calendar year. 3. *Wave of peak excess deaths per year:* We created a binary indicator for each wave indicating whether the peak excess for a state-HRR death occurred during the wave and interacted this indicator with the contemporaneous calendar year.	Beneficiaries in regions with greater excess deaths likely experienced higher severity of COVID-19 illness, reflecting poorer underlying health status. These beneficiaries are likely to have higher health care utilization and expenditures for COVID-19—in particular, higher ED and inpatient care use—than beneficiaries in regions with fewer excess deaths. However, this may be offset by more delayed or avoided care among beneficiaries in regions with greater severity of COVID-19 illness.
Pandemic Vulnerability Index (PVI)	• A measure created by the National Institute of Environmental Health Sciences, North Carolina State University, and Texas A&M University that evaluates how vulnerable a community is to COVID-19. • Using county- and state-level datasets, the PVI combines 12 indicators across four major domains: current infection rates (infection prevalence, rate of increase), baseline population concentration (daytime density/traffic, residential density), current interventions (social distancing, testing rates), and health and environmental vulnerabilities (susceptible populations, air pollution, age distribution, comorbidities, health disparities, and hospital beds). These 12 indicators are then integrated at the county level into an overall PVI score. **Year(s) available:** 2020 and 2021 **Frequency:** Monthly **Geographic level for which the variable is defined:** County **Specific variable definitions included in model testing:** Calculated the average monthly value of the PVI for the county during each wave and interacted the average values with the contemporaneous year indicator for 2020 or 2021.	Some regions may have greater vulnerability to the pandemic and higher PVI scores—for example, if they have a more susceptible population or lax local interventions. Beneficiaries in these regions are more likely to incur higher utilization and expenses related to COVID-19 than beneficiaries in regions with lower PVI scores. At the same time, regions with higher PVI scores could also experience greater indirect effects in the form of health care avoidance.

Table 5.D.5. ***(continued)***

Control	Description	Rationale for Inclusion
Government Response Index (GRI)	• The Oxford COVID-19 Government Response Tracker collects systematic information on policy measures that governments have taken to tackle COVID-19. • Policy responses are coded into 23 indicators, such as school closures, travel restrictions and vaccination policy. The GRI is a composite measure based on all 23 indicators tracked by the project. **Year(s) available:** 2020 and 2021 **Frequency:** Yearly **Geographic level for which the variable is defined:** State **Specific variable definitions included in model testing:** Interacted the values of the GRI with the contemporaneous year indicator for 2020 or 2021.	Beneficiaries in regions that had a stronger government response to COVID-19 tended to have a lower incidence of COVID-19 (Islam et al. 2020) in the long run, which may lead to lower health care utilization and expenditures for COVID-19 in the region.
Social Vulnerability Index (SVI)	• Social vulnerability refers to the resilience of communities when responding to or recovering from threats to public health. • The SVI, prepared by the Centers for Disease Control and Prevention, draws together 16 different measures of vulnerability in three themes: (1) socioeconomic (for example, poverty, unemployment), (2) demographic (for example, number of elderly and disabled) and (3) housing/transportation (for example, percentage of mobile homes, households with no vehicle). • For every measure, census tracts above the 90th percentile, or the most vulnerable 10 percent of communities, are assigned a flag. The SVI is created by counting the total number of flags in each census tract. The higher the count, the more vulnerable the population. **Year(s) available:** 2000, 2010, 2014, 2016, 2018 (only 2018 used) **Frequency:** Biannually **Geographic level for which the variable is defined:** Census tract **Specific variable definitions included in model testing:** Interacted the values of the SVI with year indicators for 2020 and 2021, separately.	Communities with higher SVI scores had higher rates of COVID-19 infections and deaths compared to communities with lower SVI scores (Freese et al. 2021; Islam et al. 2021; Karaye et al. 2020), which could have led to higher health care utilization and expenditures for COVID-19. Compared to the PVI, the SVI captures different aspects of vulnerability. For example, the SVI includes community levels of poverty, whereas the PVI focuses on COVID-19-specific measures. Also, the SVI is measured at a more granular level—census tract as opposed to county.

In addition to the COVID-19 controls presented in Table 5.D.5, we considered including controls for regional COVID-19 vaccination rates among the 65+ population (see Figure 5.D.vi in Supplement 6 for the regional distribution of COVID-19 vaccination rates within CPC+ and comparison regions). If there were differences in vaccination take-up between CPC+ and comparison regions, this could create differential changes in utilization between baseline and 2021 (the year when vaccinations against COVID-19 became available). However, CPC+ could potentially affect beneficiaries' likelihood of vaccination (either through patient education or direct administration of the vaccine in a doctor's office), and therefore vaccination rates could be endogenous. We considered including controls that capture vaccine hesitancy rates in a region (for example, flu vaccination rates, mask-wearing behavior in 2020, and the HHS/ASPE measure of vaccine hesitancy). However, similar to the concern about the COVID-19 vaccination rates, CPC+ could potentially affect these measures of vaccine hesitancy as well. Therefore, we chose not to include any of these measures as control variables. Nevertheless, our additional analyses using the actual county-level COVID-19 vaccination rates suggest that including controls for regional vaccination rates is unlikely to change our main takeaways. First, although there were regional differences in the COVID-19 vaccination rates in July 2021, the differences in CPC+ and comparison regions disappeared by the end of 2021 (Figure 5.D.2). Second, our sensitivity tests showed that including control variables for the actual county-level vaccination rates against COVID-19 midway through and at the end of 2021 did not change our main takeaways (Supplement 7, Section D).

Figure 5.D.2. Differences across the four research groups (CPC+, comparison, non-CPC+, and non-comparison) in county-level COVID-19 vaccination rates by end of 2021 were less than 1 percentage point



Source: Mathematica's analysis of CDC county-level COVID-19 vaccination rates (July 2021 and December 2021) combined with data on the number of Medicare FFS beneficiaries assigned to selected and unselected practices in 2021 (from Medicare claims and Medicare Enrollment Database).

Note: These panels show the average rates of COVID-19 vaccination in the counties of CPC+ and comparison practices as well as in the counties of non-CPC+ and non-comparison practices, by track. We show the vaccination rates for the middle of 2021 (end of July 2021) and by the end of 2021 (end of December 2021). We used CDC county-level COVID-19 vaccination rates defined as the percentage of the 65+ population in a practice's county who were fully vaccinated against COVID-19 (where, per CDC, individuals are fully vaccinated if they received one dose of a single-dose vaccine or two doses of an mRNA or protein-based series). For CPC+ practices, observations were weighted by the number of Medicare FFS beneficiaries assigned to the practice during the month and year. For comparison practices, the weight is a product of the number of assigned beneficiaries and the matching weight. For non-CPC+ and non-comparison practices, we used a concentration weight constructed at the state-HRR level such that non-CPC+ practices had the same level of representation (in terms of beneficiary months) as CPC+ practices in the same state and HRR and SSP group, and non-comparison practices had the same level of representation as comparison practices in the same state and HRR and SSP group. We winsorized the weights at the 99th percentile.

CDC = Centers for Disease Control and Prevention; CPC+ = Comprehensive Primary Care Plus; HRR = hospital referral region; SSP = Medicare Shared Savings Program; FFS = fee-for-service.

We also considered alternative ways to model COVID-19 regional controls. For example, we considered an alternative measure of excess deaths that excluded beneficiaries attributed to CPC+ or comparison practices from the calculation of the regional death rates, to minimize the potential of endogeneity in the event that CPC+ affected mortality rates during the pandemic. Although our key takeaways and model fit did not change with this alternative measure in the COVID-19 analyses, the estimates were more unstable, likely because we excluded 24 percent of beneficiaries (data not shown). Our clinical team thought that is it highly unlikely CPC+ would have differentially affected mortality during the pandemic because CPC+ and comparison practices provided many of the same services and reported little to no differences in their ability to care for patients during 2020 and 2021. Further, both groups received public health emergency funding and had similar rates of telehealth use. We also considered using latent class analysis to group regions with similar excess mortality rates and use that grouping as COVID-19 controls for the excess mortality measures. However, our testing indicated that these models performed worse than using the pandemic wave definitions in terms of model fit statistics (e.g., adjusted R-squared and other statistics).

B. Differences between CPC+ and comparison regions in health care utilization and expenditures, when including COVID-19-related control variables

We examined results from regression models that estimate the differential change in outcomes from 2019 to 2021 between (1) unselected practices in CPC+ and comparison regions, and (2) CPC+ and matched comparison practices—with and without the COVID-19 controls. We first included beneficiary controls and practice fixed effects in the models (to mimic our annual impact models as much as possible). We then compared results of these models to those where we additionally included the COVID-19-related controls. This practice-level analysis was based on outcomes aggregated at the practice level for beneficiaries who were enrolled in Medicare FFS during any given month in 2019, 2020, or 2021, and assigned to one of the four practice types: CPC+ practices, comparison practices, non-CPC+ practices, and non-comparison practices. Although non-CPC+ and non-comparison practices were not matched at baseline and thus may have had some pre-existing differences in 2019, we focus our analysis largely on the changes between these practice groups from 2019 through 2021 to estimate regional differences that are not due to the CPC+ model. A detailed description of our methods is available in Supplement 5.

If COVID-19-related controls fully account for regional differences due to the pandemic, then the adjusted differences could decrease, increase, or have no effect when we add the COVID-19-related controls to our models. For example, including the COVID-19-related controls would shrink the regional differences in cases where the pandemic exacerbated differences between CPC+ and comparison regions. Alternatively, if the pandemic reduced differences that would otherwise exist between CPC+ and comparison regions, including the COVID-19-related controls could increase the regional differences. Finally, if the pandemic did not affect regional differences, then we would expect there to be little effect on the estimated differences when comparing models with and without COVID-19-related controls. The effect of the COVID-19 controls could consequently vary by outcome and by pandemic year. As a reminder, in our AR4 COVID analysis, we found that the COVID-19-related controls generally shrank or had no effect on the regional differences in the three outcomes we studied for 2020 (Medicare Part A and B expenditures, all-cause hospitalizations, and all-cause outpatient ED visits).

Results from our current analysis showed that:

- Regional differences in the changes in outcomes remained after adjusting for beneficiary controls and practice fixed effects, but the differences were small (less than 2 percent of the 2019 mean) (Tables 5.D.6 and 5.D.7). Over the course of the pandemic (2020 and 2021 combined), the regional differences were less than 1 percent of the 2019 mean. When looking at individual pandemic years, there was more fluctuation in the regional differences, but differences were still less than 2 percent of the 2019 mean for most outcomes. The one exception, as we saw in unadjusted differences, was UCC visits, where there were larger regional differences of up to 16 percent for Track 1 and 12 percent for Track 2.
- Over the course of the pandemic (2020 and 2021 combined), including the COVID-19-related controls reduced or had little effect on the small differences between unselected practices in CPC+ and comparison regions (Tables 5.D.6 and 5.D.7). One exception is expenditures on acute hospitalizations, where including the COVID-19-related controls increased the difference from -0.01 to 0.4 percent in Track 1.
- The COVID-19-related controls had different effects for 2020 versus 2021 (Tables 5.D.6 and 5.D.7):
  - For 2020, including the COVID-19-related controls reduced or did not impact the small, adjusted differences between unselected practices in CPC+ and comparison regions.
    - **Medicare expenditures.** Consistent with AR4, the COVID-19-related controls reduced the adjusted differences in Medicare expenditures by about half. After including COVID-19-related control variables, both estimates were no longer statistically significant from zero with adjusted differences for Medicare expenditures in 2020 decreasing from -1.0 to -0.5 percent in Track 1 and from -0.8 to -0.5 percent in Track 2.
  - **Acute care use.** Adjusted differences in rates of acute hospitalizations and outpatient ED visits remained similar with the inclusion of COVID-19-related controls. Including the COVID-19-related controls also had little effect on the differences in rates of UCC visits in 2020.
    - **Acute care expenditures.** Adjusted differences for expenditures for acute hospitalizations decreased from -1.4 to -1.1 percent in Track 1 and remained statistically insignificant in Track 2 after including the COVID-19-related controls. Adjusted differences for expenditures on outpatient ED visits remained similar with the inclusion of COVID-19-related controls.
  - For 2021, including the COVID-19-related controls had less of an effect on the small differences between unselected CPC+ and comparison practices; when the controls did affect differences, results were mixed.
    - **Medicare expenditures.** For Track 1, the adjusted difference for Medicare expenditures in 2021 increased from -0.7 to -0.8 percent when adding the COVID-19-related controls, and both estimates were statistically significant. For Track 2, the difference in Medicare expenditures decreased slightly from -0.8 to -0.7 percent and the estimate was no longer statistically distinguishable from zero after including the COVID-19-related controls.
  - **Acute care use.** Including the COVID-19-related controls reduced the regional difference for the rate of acute hospitalizations from 0.3 to 0.1 percent (neither was statistically significant) in Track 1, while the COVID-19-related controls had no effect on the difference in Track 2. For outpatient ED visits, the adjusted difference decreased from 1.1 to 0.6 percent in Track 1, while it increased from -0.8 to -1.0 percent in Track 2. For rates of UCC visits in 2021, the COVID-19-related controls reduced the differences by about half, but large differences remained (over 6 percent) and were statistically significant.

    - **Acute care expenditures.** For expenditures on acute hospitalizations, including the COVID-19-related controls had little effect on the difference in Track 1, while the COVID-19-related controls increased the difference from 0.9 to 1.4 percent in Track 2. Adjusted differences for expenditures on outpatient ED visits increased from -0.6 to -1.2 percent (though neither was statistically distinguishable from zero) in Track 1 and increased from -1.4 to -1.8 percent in Track 2 (both statistically significant).

- When stratifying by SSP status, there was larger variability in the estimated differences across outcomes, particularly among SSP practices, but regional differences were still typically less than 3 percent of the 2019 mean (Tables 5.D.6 and 5.D.7). In most cases, the COVID-19-related controls reduced the differences for Track 1 in 2020 and 2021. The pattern in Track 2 was similar, except that there were more occasions where the COVID-19-related controls increased the difference in 2021. Still, regional differences after including the COVID-19-related controls remained below 3 percent, with the exception of outpatient ED expenditures (differences up to 7 percent among SSP practices in Track 2) and UCC visits (differences up to 13 percent among non-SSP practices in Track 2).

The findings for CPC+ and comparison practices were largely consistent with the findings for unselected practices in CPC+ and comparison regions, in which the COVID-19-related controls typically reduced or had little effect on the differences in 2020 but had mixed effects for 2021 (Tables 5.D.v and 5.D.vi in Supplement 7).

Table 5.D.6. Using COVID-19-related controls reduced or had little effect on the regional differences among unselected practices for 2020, but the pattern was mixed for 2021, particularly among SSP practices (Track 1)

	Track 1 – all, Unadjusted means		Track 1 – all, Regression-adjusted non-CPC+ vs. non-comparison differences relative to 2019		Track 1 – SSP, Unadjusted means		Track 1 – SSP, Regression-adjusted non-CPC+ vs. non-comparison differences relative to 2019		Track 1 – non-SSP, Unadjusted means		Track 1 – non-SSP, Regression-adjusted non-CPC+ vs. non-comparison differences relative to 2019	
	Non-CPC+	Non-comparison	Estimate (SE)	Percentage difference relative to 2019 non-comparison mean[a]	Non-CPC+	Non-comparison	Estimate (SE)	Percentage difference relative to 2019 non-comparison mean[a]	Non-CPC+	Non-comparison	Estimate (SE)	Percentage difference relative to 2019 non-comparison mean[a]
Medicare Part A and B expenditures without enhanced payments for CPC+ and SSP												
Using practice fixed effects and beneficiary controls												
2019 (Jan–Dec)	$1,045	$1,059	NA	NA	$1,066	$1,080	NA	NA	$1,023	$1,034	NA	NA
2020 (Jan–Dec)	$994	$1,019	-$10.4*** ($3.7)	-1.0%	$1,015	$1,032	-$5.8 ($5.9)	-0.5%	$973	$1,002	-$15.2*** ($4.5)	-1.5%
2021 (Jan–Dec)	$1,071	$1,097	-$7.4* ($4.3)	-0.7%	$1,100	$1,110	$4.6 ($7.0)	0.4%	$1,040	$1,080	-$19.9*** ($5.1)	-1.9%
2020–2021 (combined COVID-19 years)	$1,032	$1,058	-$8.9** ($3.5)	-0.8%	$1,056	$1,070	-$0.7 ($5.6)	-0.1%	$1,006	$1,040	-$17.5*** ($4.2)	-1.7%
Using practice fixed effects and beneficiary controls, and COVID-19-related controls												
2019 (Jan–Dec)	$1,045	$1,059	NA	NA	$1,066	$1,080	NA	NA	$1,023	$1,034	NA	NA
2020 (Jan–Dec)	$994	$1,019	-$5.5 ($4.2)	-0.5%	$1,015	$1,032	-$1.0 ($6.7)	-0.1%	$973	$1,002	-$12.2** ($5.0)	-1.2%
2021 (Jan–Dec)	$1,071	$1,097	-$8.0* ($4.5)	-0.8%	$1,100	$1,110	$1.2 ($7.7)	0.1%	$1,040	$1,080	-$19.2*** ($5.4)	-1.9%
2020–2021 (combined COVID-19 years)	$1,032	$1,058	-$6.7* ($3.7)	-0.6%	$1,056	$1,070	$0.0 ($6.1)	0.00%	$1,006	$1,040	-$15.8*** ($4.4)	-1.5%

Table 5.D.6. ***(continued)***

	Track 1 – all, Unadjusted means		Track 1 – all, Regression-adjusted non-CPC+ vs. non-comparison differences relative to 2019		Track 1 – SSP, Unadjusted means		Track 1 – SSP, Regression-adjusted non-CPC+ vs. non-comparison differences relative to 2019		Track 1 – non-SSP, Unadjusted means		Track 1 – non-SSP, Regression-adjusted non-CPC+ vs. non-comparison differences relative to 2019	
	Non-CPC+	Non-comparison	Estimate (SE)	Percentage difference relative to 2019 non-comparison mean[a]	Non-CPC+	Non-comparison	Estimate (SE)	Percentage difference relative to 2019 non-comparison mean[a]	Non-CPC+	Non-comparison	Estimate (SE)	Percentage difference relative to 2019 non-comparison mean[a]
Acute hospitalizations (short-stay acute care and critical access hospitals)												
Using practice fixed effects and beneficiary controls												
2019 (Jan–Dec)	312	303	NA	NA	315	303	NA	NA	310	302	NA	NA
2020 (Jan–Dec)	264	259	-2.9** (1.4)	-1.0%	267	260	-2.4 (2.3)	-0.8%	260	259	-3.5** (1.7)	-1.2%
2021 (Jan–Dec)	262	255	1.0 (1.6)	0.3%	267	255	4.0 (2.6)	1.3%	256	255	-2.1 (1.9)	-0.7%
2020–2021 (combined COVID-19 years)	263	257	-1.0 (1.4)	-0.3%	267	257	0.7 (2.2)	0.2%	258	257	-2.8* (1.6)	-0.9%
Using practice fixed effects and beneficiary controls, and COVID-19-related controls												
2019 (Jan–Dec)	312	303	NA	NA	315	303	NA	NA	310	302	NA	NA
2020 (Jan–Dec)	264	259	-2.9* (1.7)	-1.0%	267	260	-2.7 (2.7)	-0.9%	260	259	-3.6* (1.9)	-1.2%
2021 (Jan–Dec)	262	255	0.3 (1.7)	0.1%	267	255	3.6 (3.0)	1.2%	256	255	-2.7 (1.9)	-0.9%
2020–2021 (combined COVID-19 years)	263	257	-1.3 (1.5)	-0.4%	267	257	0.3 (2.5)	0.1%	258	257	-3.1* (1.6)	-1.0%
Expenditures on acute hospitalizations (short-stay acute care and critical access hospitals) (per beneficiary per month)												
Using practice fixed effects and beneficiary controls												
2019 (Jan–Dec)	$327	$332	NA	NA	$333	$337	NA	NA	$320	$325	NA	NA
2020 (Jan–Dec)	$307	$318	-$4.6** ($2.2)	-1.4%	$314	$322	-$4.1 ($3.4)	-1.2%	$301	$313	-$5.2* ($2.7)	-1.6%
2021 (Jan–Dec)	$319	$327	-$0.9 ($2.3)	-0.3%	$327	$328	$3.8 ($3.6)	1.1%	$309	$324	-$6.0** ($2.7)	-1.9%
2020–2021 (combined COVID-19 years)	$313	$322	-$2.8 ($1.9)	-0.8%	$321	$325	-$0.2 ($3.0)	-0.1%	$305	$318	-$5.6** ($2.3)	-1.7%

Table 5.D.6. ***(continued)***

	Track 1 – all, Unadjusted means		Track 1 – all, Regression-adjusted non-CPC+ vs. non-comparison differences relative to 2019		Track 1 – SSP, Unadjusted means		Track 1 – SSP, Regression-adjusted non-CPC+ vs. non-comparison differences relative to 2019		Track 1 – non-SSP, Unadjusted means		Track 1 – non-SSP, Regression-adjusted non-CPC+ vs. non-comparison differences relative to 2019	
	Non-CPC+	Non-comparison	Estimate (SE)	Percentage difference relative to 2019 non-comparison mean[a]	Non-CPC+	Non-comparison	Estimate (SE)	Percentage difference relative to 2019 non-comparison mean[a]	Non-CPC+	Non-comparison	Estimate (SE)	Percentage difference relative to 2019 non-comparison mean[a]
Using practice fixed effects and beneficiary controls, and COVID-19-related controls												
2019 (Jan–Dec)	$327	$332	NA	NA	$333	$337	NA	NA	$320	$325	NA	NA
2020 (Jan–Dec)	$307	$318	-$3.7 ($2.5)	-1.1%	$314	$322	-$4.9 ($3.9)	-1.5%	$301	$313	-$3.6 ($3.0)	-1.1%
2021 (Jan–Dec)	$319	$327	-$0.7 ($2.4)	-0.2%	$327	$328	$3.6 ($4.0)	1.1%	$309	$324	-$5.6* ($2.9)	-1.7%
2020–2021 (combined COVID-19 years)	$313	$322	-$2.2 ($2.1)	-0.7%	$321	$325	-$0.9 ($3.3)	-0.3%	$305	$318	-$4.7* ($2.4)	-1.4%
Outpatient ED visits, including observation stays												
Using practice fixed effects and beneficiary controls												
2019 (Jan–Dec)	524	531	NA	NA	509	509	NA	NA	543	550	NA	NA
2020 (Jan–Dec)	404	416	2.0 (2.3)	0.4%	388	398	-0.3 (4.1)	-0.1%	424	433	4.1 (2.7)	0.7%
2021 (Jan–Dec)	431	442	5.9** (2.6)	1.1%	418	424	6.7* (4.0)	1.3%	448	460	4.3 (3.2)	0.8%
2020–2021 (combined COVID-19 years)	417	429	3.9* (2.2)	0.7%	402	411	3.1 (3.6)	0.6%	436	446	4.2 (2.7)	0.8%
Using practice fixed effects and beneficiary controls, and COVID-19-related controls												
2019 (Jan–Dec)	524	531	NA	NA	509	509	NA	NA	543	550	NA	NA
2020 (Jan–Dec)	404	416	2.1 (2.6)	0.4%	388	398	2.0 (4.8)	0.4%	424	433	2.2 (2.9)	0.4%
2021 (Jan–Dec)	431	442	3.0 (2.8)	0.6%	418	424	4.0 (4.5)	0.8%	448	460	1.2 (3.4)	0.2%
2020–2021 (combined COVID-19 years)	417	429	2.5 (2.4)	0.5%	402	411	2.9 (4.1)	0.6%	436	446	1.7 (2.8)	0.3%

Table 5.D.6. ***(continued)***

	Track 1 – all, Unadjusted means		Track 1 – all, Regression-adjusted non-CPC+ vs. non-comparison differences relative to 2019		Track 1 – SSP, Unadjusted means		Track 1 – SSP, Regression-adjusted non-CPC+ vs. non-comparison differences relative to 2019		Track 1 – non-SSP, Unadjusted means		Track 1 – non-SSP, Regression-adjusted non-CPC+ vs. non-comparison differences relative to 2019	
	Non-CPC+	Non-comparison	Estimate (SE)	Percentage difference relative to 2019 non-comparison mean[a]	Non-CPC+	Non-comparison	Estimate (SE)	Percentage difference relative to 2019 non-comparison mean[a]	Non-CPC+	Non-comparison	Estimate (SE)	Percentage difference relative to 2019 non-comparison mean[a]
Expenditures on outpatient ED visits, including observation stays (per beneficiary per month)												
Using practice fixed effects and beneficiary controls												
2019 (Jan–Dec)	$31	$33	NA	NA	$30	$31	NA	NA	$32	$34	NA	NA
2020 (Jan–Dec)	$26	$28	-$0.1 ($0.2)	-0.3%	$24	$26	-$0.4 ($0.3)	-1.2%	$28	$29	$0.2 ($0.2)	0.6%
2021 (Jan–Dec)	$30	$32	-$0.2 ($0.3)	-0.6%	$28	$31	-$1.0** ($0.4)	-3.2%	$32	$33	$0.5 ($0.4)	1.4%
2020–2021 (combined COVID-19 years)	$28	$30	-$0.1 ($0.2)	-0.4%	$26	$28	-$0.7** ($0.3)	-2.2%	$30	$31	$0.3 ($0.3)	1.0%
Using practice fixed effects and beneficiary controls, and COVID-19-related controls												
2019 (Jan–Dec)	$31	$33	NA	NA	$30	$31	NA	NA	$32	$34	NA	NA
2020 (Jan–Dec)	$26	$28	$0.1 ($0.2)	0.4%	$24	$26	$0.0 ($0.4)	-0.1%	$28	$29	$0.3 ($0.3)	0.8%
2021 (Jan–Dec)	$30	$32	-$0.4 ($0.3)	-1.2%	$28	$31	-$1.4*** ($0.5)	-4.6%	$32	$33	$0.2 ($0.4)	0.5%
2020–2021 (combined COVID-19 years)	$28	$30	-$0.1 ($0.2)	-0.4%	$26	$28	-$0.7** ($0.3)	-2.2%	$30	$31	$0.2 ($0.3)	0.7%
Urgent care center visits (per 1,000 beneficiaries per year)												
Using practice fixed effects and beneficiary controls												
2019 (Jan–Dec)	124	129	NA	NA	136	140	NA	NA	112	119	NA	NA
2020 (Jan–Dec)	135	128	13.4*** (1.6)	10.4%	151	139	18.1*** (2.7)	12.9%	118	116	9.8*** (1.8)	8.3%
2021 (Jan–Dec)	188	175	20.2*** (2.6)	15.6%	213	195	26.1*** (4.5)	18.6%	160	154	14.5*** (2.8)	12.2%
2020–2021 (combined COVID-19 years)	161	151	16.7*** (2.0)	12.9%	181	166	22.0*** (3.3)	15.7%	139	135	12.1*** (2.1)	10.2%

Table 5.D.6. ***(continued)***

	Track 1 – all, Unadjusted means		Track 1 – all, Regression-adjusted non-CPC+ vs. non-comparison differences relative to 2019		Track 1 – SSP, Unadjusted means		Track 1 – SSP, Regression-adjusted non-CPC+ vs. non-comparison differences relative to 2019		Track 1 – non-SSP, Unadjusted means		Track 1 – non-SSP, Regression-adjusted non-CPC+ vs. non-comparison differences relative to 2019	
	Non-CPC+	Non-comparison	Estimate (SE)	Percentage difference relative to 2019 non-comparison mean[a]	Non-CPC+	Non-comparison	Estimate (SE)	Percentage difference relative to 2019 non-comparison mean[a]	Non-CPC+	Non-comparison	Estimate (SE)	Percentage difference relative to 2019 non-comparison mean[a]
Using practice fixed effects and beneficiary controls, and COVID-19-related controls												
2019 (Jan–Dec)	124	129	NA	NA	136	140	NA	NA	112	119	NA	NA
2020 (Jan–Dec)	135	128	13.0*** (1.8)	10.1%	151	139	13.5*** (3.0)	9.6%	118	116	12.3*** (1.9)	10.3%
2021 (Jan–Dec)	188	175	7.9*** (2.6)	6.1%	213	195	-1.3 (4.9)	-1.0%	160	154	12.9*** (2.7)	10.8%
2020–2021 (combined COVID-19 years)	161	151	10.5*** (2.0)	8.1%	181	166	6.5* (3.4)	4.6%	139	135	12.6*** (2.1)	10.6%
Unweighted sample sizes												
Number of practices	8,337	20,654			2,488	5,151			5,849	15,503		
Average number of beneficiaries per month	2,284,100	6,798,349			809,545	2,167,584			1,474,555	4,630,765		

Source: Mathematica's analysis of Medicare claims data from January 2019 through December 2021.

Note: Estimates in the table are derived from separate models run at the practice-month-year level that are regression-adjusted for (1) baseline beneficiary characteristics and practice fixed effects and (2) baseline beneficiary characteristics, practice fixed effects, and COVID-19-related controls. For these practices, we used a concentration weight constructed at the state-HRR level such that non-CPC+ practices had the same level of representation (in terms of beneficiary months) as CPC+ practices in the same state and HRR and SSP group, and non-comparison practices had the same level of representation as comparison practices in the same state and HRR and SSP group. We winsorized the weights at the 99th percentile. Standard errors are clustered at the practice level. For a detailed description of methods, see Supplement 5.

*/**/*** Significantly different from zero at the 0.10/0.05/0.01 levels, two-tailed test.

[a] To calculate these percentages, we divided the non-CPC+ vs. non-comparison differences by the unadjusted 2019 non-comparison mean for the outcome.

CPC+ = Comprehensive Primary Care Plus; ED = emergency department; FFS = fee-for-service; HRR = hospital referral region; NA = not applicable; SE = standard error; SSP = Medicare Shared Savings Program.

Table 5.D.7. Using COVID-19-related controls reduced or had little effect on the regional differences among unselected practices for 2020, but the pattern was mixed for 2021, particularly among SSP practices (Track 2)

	Track 2 – all, Unadjusted means		Track 2 – all, Regression-adjusted non-CPC+ vs. non-comparison differences relative to 2019		Track 2 – SSP, Unadjusted means		Track 2 – SSP, Regression-adjusted non-CPC+ vs. non-comparison differences relative to 2019		Track 2 – non-SSP, Unadjusted means		Track 2 – non-SSP, Regression-adjusted non-CPC+ vs. non-comparison differences relative to 2019	
	Non-CPC+	Non-comparison	Estimate (SE)	Percentage difference relative to 2019 non-comparison mean[a]	Non-CPC+	Non-comparison	Estimate (SE)	Percentage difference relative to 2019 non-comparison mean[a]	Non-CPC+	Non-comparison	Estimate (SE)	Percentage difference relative to 2019 non-comparison mean[a]
Medicare Part A and B expenditures without enhanced payments for CPC+ and SSP												
Using practice fixed effects and beneficiary controls												
2019 (Jan–Dec)	$1,036	$1,057	NA	NA	$1,063	$1,087	NA	NA	$1,015	$1,026	NA	NA
2020 (Jan–Dec)	$986	$1,015	-$8.4** ($3.7)	-0.8%	$1,006	$1,039	-$11.4* ($6.4)	-1.1%	$970	$991	-$7.3* ($4.4)	-0.7%
2021 (Jan–Dec)	$1,058	$1,091	-$8.4* ($4.3)	-0.8%	$1,084	$1,113	-$4.9 ($7.7)	-0.5%	$1,036	$1,068	-$13.2*** ($5.1)	-1.3%
2020–2021 (combined COVID-19 years)	$1,021	$1,053	-$8.4** ($3.5)	-0.8%	$1,045	$1,075	-$8.2 ($6.1)	-0.8%	$1,002	$1,028	-$10.2** ($4.2)	-1.0%
Using practice fixed effects and beneficiary controls, and COVID-19-related controls												
2019 (Jan–Dec)	$1,036	$1,057	NA	NA	$1,063	$1,087	NA	NA	$1,015	$1,026	NA	NA
2020 (Jan–Dec)	$986	$1,015	-$5.7 ($4.2)	-0.5%	$1,006	$1,039	-$7.5 ($7.1)	-0.7%	$970	$991	-$5.5 ($4.9)	-0.5%
2021 (Jan–Dec)	$1,058	$1,091	-$7.4 ($4.5)	-0.7%	$1,084	$1,113	-$4.5 ($8.7)	-0.4%	$1,036	$1,068	-$10.8** ($5.4)	-1.1%
2020–2021 (combined COVID-19 years)	$1,021	$1,053	-$6.5* ($3.8)	-0.6%	$1,045	$1,075	-$6.0 ($6.8)	-0.6%	$1,002	$1,028	-$8.1* ($4.5)	-0.8%
Acute hospitalizations (short-stay acute care and critical access hospitals)												
Using practice fixed effects and beneficiary controls												
2019 (Jan–Dec)	312	306	NA	NA	327	309	NA	NA	301	302	NA	NA
2020 (Jan–Dec)	265	261	-1.3 (1.4)	-0.4%	279	265	-2.9 (2.6)	-1.0%	255	258	-0.7 (1.6)	-0.2%
2021 (Jan–Dec)	262	256	2.5 (1.6)	0.8%	279	258	5.4* (3.0)	1.7%	250	254	-0.4 (1.8)	-0.1%
2020–2021 (combined COVID-19 years)	264	259	0.5 (1.4)	0.2%	279	261	1.1 (2.5)	0.4%	252	256	-0.6 (1.5)	-0.2%

Table 5.D.7. ***(continued)***

	Track 2 – all, Unadjusted means		Track 2 – all, Regression-adjusted non-CPC+ vs. non-comparison differences relative to 2019		Track 2 – SSP, Unadjusted means		Track 2 – SSP, Regression-adjusted non-CPC+ vs. non-comparison differences relative to 2019		Track 2 – non-SSP, Unadjusted means		Track 2 – non-SSP, Regression-adjusted non-CPC+ vs. non-comparison differences relative to 2019	
	Non-CPC+	Non-comparison	Estimate (SE)	Percentage difference relative to 2019 non-comparison mean[a]	Non-CPC+	Non-comparison	Estimate (SE)	Percentage difference relative to 2019 non-comparison mean[a]	Non-CPC+	Non-comparison	Estimate (SE)	Percentage difference relative to 2019 non-comparison mean[a]
Using practice fixed effects and beneficiary controls, and COVID-19-related controls												
2019 (Jan–Dec)	312	306	NA	NA	327	309	NA	NA	301	302	NA	NA
2020 (Jan–Dec)	265	261	-1.7 (1.6)	-0.6%	279	265	-3.6 (2.9)	-1.2%	255	258	-0.7 (1.8)	-0.2%
2021 (Jan–Dec)	262	256	2.4 (1.7)	0.8%	279	258	5.8* (3.4)	1.9%	250	254	-0.2 (1.9)	-0.1%
2020–2021 (combined COVID-19 years)	264	259	0.4 (1.5)	0.1%	279	261	1.0 (2.7)	0.3%	252	256	-0.4 (1.6)	-0.2%
Expenditures on acute hospitalizations (short-stay acute care and critical access hospitals) (per beneficiary per month)												
Using practice fixed effects and beneficiary controls												
2019 (Jan–Dec)	$327	$334	NA	NA	$334	$342	NA	NA	$321	$324	NA	NA
2020 (Jan–Dec)	$308	$319	-$3.0 ($2.2)	-0.9%	$316	$327	-$3.0 ($3.6)	-0.9%	$302	$309	-$2.8 ($2.7)	-0.9%
2021 (Jan–Dec)	$320	$326	$3.1 ($2.4)	0.9%	$332	$332	$9.3** ($4.0)	2.7%	$311	$319	-$2.6 ($2.9)	-0.8%
2020–2021 (combined COVID-19 years)	$314	$322	$0.0 ($2.0)	-0.01%	$324	$329	$3.0 ($3.3)	0.9%	$307	$314	-$2.7 ($2.4)	-0.8%
Using practice fixed effects and beneficiary controls, and COVID-19-related controls												
2019 (Jan–Dec)	$327	$334	NA	NA	$334	$342	NA	NA	$321	$324	NA	NA
2020 (Jan–Dec)	$308	$319	-$2.1 ($2.4)	-0.6%	$316	$327	-$2.2 ($4.0)	-0.7%	$302	$309	-$1.7 ($2.9)	-0.5%
2021 (Jan–Dec)	$320	$326	$4.7* ($2.5)	1.4%	$332	$332	$10.3** ($4.5)	3.0%	$311	$319	-$1.1 ($3.1)	-0.3%
2020–2021 (combined COVID-19 years)	$314	$322	$1.3 ($2.1)	0.4%	$324	$329	$3.9 ($3.6)	1.1%	$307	$314	-$1.4 ($2.6)	-0.4%

Table 5.D.7. ***(continued)***

	Track 2 – all, Unadjusted means		Track 2 – all, Regression-adjusted non-CPC+ vs. non-comparison differences relative to 2019		Track 2 – SSP, Unadjusted means		Track 2 – SSP, Regression-adjusted non-CPC+ vs. non-comparison differences relative to 2019		Track 2 – non-SSP, Unadjusted means		Track 2 – non-SSP, Regression-adjusted non-CPC+ vs. non-comparison differences relative to 2019	
	Non-CPC+	Non-comparison	Estimate (SE)	Percentage difference relative to 2019 non-comparison mean[a]	Non-CPC+	Non-comparison	Estimate (SE)	Percentage difference relative to 2019 non-comparison mean[a]	Non-CPC+	Non-comparison	Estimate (SE)	Percentage difference relative to 2019 non-comparison mean[a]
Outpatient ED visits, including observation stays												
Using practice fixed effects and beneficiary controls												
2019 (Jan–Dec)	547	534	NA	NA	522	518	NA	NA	567	548	NA	NA
2020 (Jan–Dec)	421	419	-5.1** (2.3)	-1.0%	397	408	-11.4*** (4.4)	-2.2%	441	431	-1.9 (2.7)	-0.4%
2021 (Jan–Dec)	446	447	-4.1 (2.8)	-0.8%	424	434	-9.4* (5.0)	-1.8%	464	458	-1.8 (3.3)	-0.3%
2020–2021 (combined COVID-19 years)	434	433	-4.6** (2.3)	-0.9%	410	420	-10.4** (4.2)	-2.0%	453	444	-1.9 (2.7)	-0.3%
Using practice fixed effects and beneficiary controls, and COVID-19-related controls												
2019 (Jan–Dec)	547	534	NA	NA	522	518	NA	NA	567	548	NA	NA
2020 (Jan–Dec)	421	419	-6.1** (2.5)	-1.2%	397	408	-10.8** (4.6)	-2.1%	441	431	-3.3 (2.9)	-0.6%
2021 (Jan–Dec)	446	447	-5.3* (3.0)	-1.0%	424	434	-13.0** (5.6)	-2.5%	464	458	-0.9 (3.3)	-0.2%
2020–2021 (combined COVID-19 years)	434	433	-5.7** (2.4)	-1.1%	410	420	-11.9*** (4.4)	-2.3%	453	444	-2.1 (2.8)	-0.4%
Expenditures on outpatient ED visits, including observation stays (per beneficiary per month)												
Using practice fixed effects and beneficiary controls												
2019 (Jan–Dec)	$32	$33	NA	NA	$31	$33	NA	NA	$33	$33	NA	NA
2020 (Jan–Dec)	$27	$28	-$0.5*** ($0.2)	-1.6%	$25	$28	-$1.3*** ($0.4)	-3.8%	$28	$28	-$0.1 ($0.2)	-0.3%
2021 (Jan–Dec)	$31	$32	-$0.5* ($0.3)	-1.4%	$29	$32	-$1.9*** ($0.5)	-5.9%	$33	$32	$0.3 ($0.3)	0.8%
2020–2021 (combined COVID-19 years)	$29	$30	-$0.5** ($0.2)	-1.5%	$27	$30	-$1.6*** ($0.4)	-4.8%	$30	$30	$0.1 ($0.3)	0.3%

Table 5.D.7. ***(continued)***

	Track 2 – all, Unadjusted means		Track 2 – all, Regression-adjusted non-CPC+ vs. non-comparison differences relative to 2019		Track 2 – SSP, Unadjusted means		Track 2 – SSP, Regression-adjusted non-CPC+ vs. non-comparison differences relative to 2019		Track 2 – non-SSP, Unadjusted means		Track 2 – non-SSP, Regression-adjusted non-CPC+ vs. non-comparison differences relative to 2019	
	Non-CPC+	Non-comparison	Estimate (SE)	Percentage difference relative to 2019 non-comparison mean[a]	Non-CPC+	Non-comparison	Estimate (SE)	Percentage difference relative to 2019 non-comparison mean[a]	Non-CPC+	Non-comparison	Estimate (SE)	Percentage difference relative to 2019 non-comparison mean[a]
Using practice fixed effects and beneficiary controls, and COVID-19-related controls												
2019 (Jan–Dec)	$32	$33	NA	NA	$31	$33	NA	NA	$33	$33	NA	NA
2020 (Jan–Dec)	$27	$28	-$0.5** ($0.2)	-1.5%	$25	$28	-$1.2*** ($0.4)	-3.7%	$28	$28	$0.0 ($0.3)	-0.1%
2021 (Jan–Dec)	$31	$32	-$0.6** ($0.3)	-1.8%	$29	$32	-$2.2*** ($0.5)	-6.6%	$33	$32	$0.1 ($0.4)	0.3%
2020–2021 (combined COVID-19 years)	$29	$30	-$0.6** ($0.2)	-1.7%	$27	$30	-$1.7*** ($0.4)	-5.1%	$30	$30	$0.0 ($0.3)	0.1%
Urgent care center visits (per 1,000 beneficiaries per year)												
Using practice fixed effects and beneficiary controls												
2019 (Jan–Dec)	123	127	NA	NA	130	135	NA	NA	118	119	NA	NA
2020 (Jan–Dec)	128	123	8.1*** (1.8)	6.4%	135	132	7.5** (3.4)	5.5%	123	115	8.4*** (2.0)	7.1%
2021 (Jan–Dec)	177	165	15.8*** (2.7)	12.5%	187	178	13.3*** (5.1)	9.8%	170	152	18.0*** (3.1)	15.1%
2020–2021 (combined COVID-19 years)	152	144	11.9*** (2.1)	9.4%	161	154	10.3*** (3.9)	7.6%	146	133	13.1*** (2.3)	11.0%
Using practice fixed effects and beneficiary controls, and COVID-19-related controls												
2019 (Jan–Dec)	123	127	NA	NA	130	135	NA	NA	118	119	NA	NA
2020 (Jan–Dec)	128	123	8.6*** (1.9)	6.8%	135	132	5.1 (3.5)	3.7%	123	115	9.9*** (2.1)	8.4%
2021 (Jan–Dec)	177	165	10.1*** (2.7)	7.9%	187	178	-1.1 (6.0)	-0.8%	170	152	15.7*** (3.1)	13.2%
2020–2021 (combined COVID-19 years)	152	144	9.4*** (2.1)	7.4%	161	154	2.0 (4.3)	1.5%	146	133	12.8*** (2.4)	10.8%

Table 5.D.7. ***(continued)***

	Track 2 – all, Unadjusted means		Track 2 – all, Regression-adjusted non-CPC+ vs. non-comparison differences relative to 2019		Track 2 – SSP, Unadjusted means		Track 2 – SSP, Regression-adjusted non-CPC+ vs. non-comparison differences relative to 2019		Track 2 – non-SSP, Unadjusted means		Track 2 – non-SSP, Regression-adjusted non-CPC+ vs. non-comparison differences relative to 2019	
	Non-CPC+	Non-comparison	Estimate (SE)	Percentage difference relative to 2019 non-comparison mean[a]	Non-CPC+	Non-comparison	Estimate (SE)	Percentage difference relative to 2019 non-comparison mean[a]	Non-CPC+	Non-comparison	Estimate (SE)	Percentage difference relative to 2019 non-comparison mean[a]
Unweighted sample sizes												
Number of practices	7,276	20,113			2,423	5,010			4,853	15,103		
Average number of beneficiaries per month	1,871,011	6,619,682			756,018	2,094,775			1,114,993	4,524,907		

Source: Mathematica's analysis of Medicare claims data from January 2019 through December 2021.

Note: Estimates in the table are derived from separate models run at the practice-month-year level that are regression-adjusted for (1) baseline beneficiary characteristics and practice fixed effects and (2) baseline beneficiary characteristics, practice fixed effects, and COVID-19-related controls. For these practices, we used a concentration weight constructed at the state-HRR level such that non-CPC+ practices had the same level of representation (in terms of beneficiary months) as CPC+ practices in the same state and HRR and SSP group, and non-comparison practices had the same level of representation as comparison practices in the same state and HRR and SSP group. We winsorized the weights at the 99th percentile. Standard errors are clustered at the practice level. For a detailed description of methods, see Supplement 5.

*/**/*** Significantly different from zero at the 0.10/0.05/0.01 levels, two-tailed test.

[a] To calculate these percentages, we divided the non-CPC+ vs. non-comparison differences by the unadjusted 2019 non-comparison mean for the outcome.

CPC+ = Comprehensive Primary Care Plus; ED = emergency department; FFS = fee-for-service; HRR = hospital referral region; NA = not applicable; SE = standard error; SSP = Medicare Shared Savings Program.

5.D.5. Key findings from the COVID-19 analyses and decisions for AR5

Below we summarize our key findings on the direct and total effects of COVID-19 on excess deaths, health care utilization, and Medicare expenditures, as well as the approach we took in the CPC+ impact evaluation in AR5 to mitigate any potential bias in the PY 4 and PY 5 impact estimates due to COVID-19.

A. Direct effects of COVID-19

- In both CPC+ and comparison regions, approximately 9 percent of beneficiaries had a COVID-19-related diagnosis in 2021, up from 7 percent in 2020. Diagnoses peaked in December 2020/January 2021 for both regions. The differences between CPC+ and comparison regions in the rate of COVID-19-related diagnoses each month was small (less than 0.1 percentage points). Compared to the first few months of the pandemic, regional differences in the rate of COVID-19-related diagnoses were smaller later in the pandemic, particularly in the second half of 2021.
- There were lower rates of excess deaths in 2021 (12 percent excess deaths) compared with 2020 (20 percent excess deaths), likely driven by the availability of COVID-19 vaccines, more effective treatment strategies, and less deadly virus variants. There was a similar timing in the peaks of excess deaths between CPC+ and comparison regions, with small differences each month (typically less than 1 death per 10,000 beneficiaries per month).

B. Total effects of COVID-19

- Over the course of the pandemic, nearly all utilization and expenditure outcomes declined from their 2019 levels, with larger declines in 2020 compared to 2021, when utilization began to return to its pre-pandemic levels. In the case of Medicare expenditures without enhanced payments, expenditures for unselected practices dipped by 4 to 5 percent of their 2019 levels in 2020, before surpassing 2019 levels by 2 to 4 percent in 2021. Unlike other outcomes, UCC visits had substantial growth over the course of the pandemic.
- Unselected practices in CPC+ regions experienced 1 to 2 percent greater reductions in health care utilization and expenditures than unselected comparison practices for most outcomes, and these regional differences were similar for 2020 and 2021. In the case of acute hospitalizations, regional differences were smaller in 2021 (less than 1 percent difference) than 2020 (less than 2 percent difference).
- With the exception of UCC visits, regional differences in the 2019 to 2020 change between unselected practices in CPC+ and comparison regions were not substantially greater in magnitude than differential year-to-year changes observed in prior years (less than 2 percent differences). Similarly, the regional differences in the changes from 2019 to 2021 were not substantially greater than historical differences in two-year–to–two-year changes (less than 3 percent differences).
- Unselected practices in CPC+ regions had substantially larger growth in UCC visits than unselected comparison practices over the pandemic, with regional differences increasing between 2020 and 2021 (up to 14 percent greater increases for unselected practices in CPC+ regions in 2021).
- When stratifying by SSP status, there were larger regional differences across most utilization and expenditure outcomes compared to overall track differences, but they still were typically less than 3 percent.

C. Accounting for potential differences due to COVID-19

- Adding the same type of COVID-19-related regional controls as in AR4 (excess deaths, indices of pandemic and social vulnerability, and indices of state policy responses)—now updated to include an additional year of data for 2021—reduced regional differences for most outcomes in 2020 but had less of a consistent effect in 2021. For example, the COVID-19 controls reduced the 1 percent difference in Medicare expenditures by almost one-half in 2020, but had almost no effect on the 1 percent difference in expenditures in 2021. The large, regional differences for UCC visits remained in both years even with the inclusion of COVID-19-related controls.
- The reason that the COVID-19-related controls seemed to matter less in 2021 could be because there were fewer systematic differences between CPC+ and comparison regions in 2021 compared to 2020. For example, the rates of COVID-19 diagnoses and acute hospitalizations were more similar in 2021 across regions, compared to 2020. Further, there was a general move toward "back to normal" in 2021, marked by lower excess death rates, the availability of at-home tests, rates of service utilization and expenditures that approached (or surpassed) their pre-pandemic levels, and similar rates of fully vaccinated individuals age 65 or older across regions.

D. Integrating our findings into the CPC+ impact evaluation for AR5

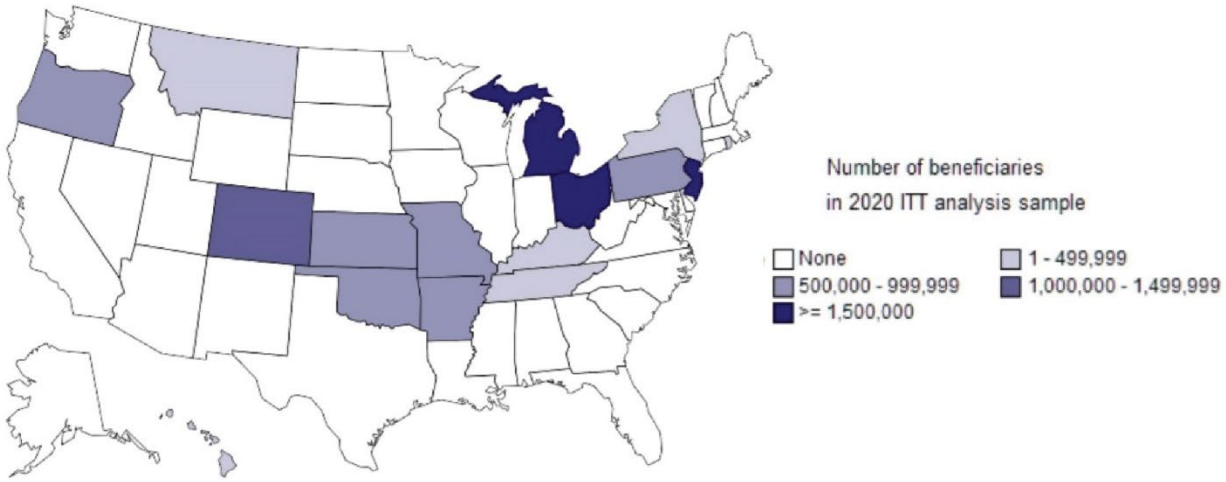
- Our findings in this Appendix indicate that the regional differences in the direct and the total effects of COVID-19 between CPC+ and comparison regions were small in 2020 and 2021. Adding regional COVID-19 controls generally reduced regional differences in 2020 but had less of a systematic effect in 2021. We believe this difference across years was because there were more health care disruptions in 2020, particularly in the first few months of the pandemic, while 2021 was marked by a return to normalcy.
- We proceeded using a difference-in-differences model with regional COVID-19-related controls as the primary analysis for the CPC+ impact evaluation in AR5, just as we did in AR4. Although the regional differences we observed were typically small, including COVID-19-related controls could help mitigate the potential bias from the pandemic in the CPC+ impact estimates for the last two program years, and for PY 4 in particular.
- In addition to using COVID-19-related controls for the CPC+ impact evaluation, we also proceeded with:
  - The triple-differences model as a key sensitivity test.
  - An additional sensitivity test for the key outcomes of Medicare expenditures without enhanced payments, number of all-cause acute hospitalizations, and number of outpatient ED visits that excludes claims during peak health care avoidance months of March through May 2020.
  - An alternative measure of UCC visits as a sensitivity test. In this alternative version, we excluded visits with claims for COVID-19-related diagnoses. This was motivated by UCC visits having different patterns of change before and during the pandemic compared to other outcomes we studied and a substantial share of UCC visits being for COVID-19-related conditions.
  - Cautious interpretation of findings, especially if the magnitude and direction of impact estimates in 2020 and 2021 differed meaningfully from those in previous years.

Supplement 1. Regional distribution of CPC+ and comparison practices

Figure 5.D.i shows the regional distribution of Track 1 and Track 2 CPC+ and comparison practices. Darker colors indicate states with a higher concentration of beneficiaries included in the 2020 ITT analysis sample. New Jersey had the largest concentration of Track 1 CPC+ beneficiaries (18 percent of all Track 1 CPC+ beneficiaries) and made up a smaller proportion of the Track 2 CPC+ ITT analysis sample (10 percent). Ohio had the largest concentration of Track 2 CPC+ beneficiaries (20 percent). Illinois and Pennsylvania had the highest concentration of comparison beneficiaries, making up 11 to 12 percent of the Track 1 and Track 2 comparison ITT sample, respectively.

Figure 5.D.i. CPC+ and comparison practices were selected from different regions[a]

Panel A. Track 1 CPC+



Panel B. Track 1 comparison

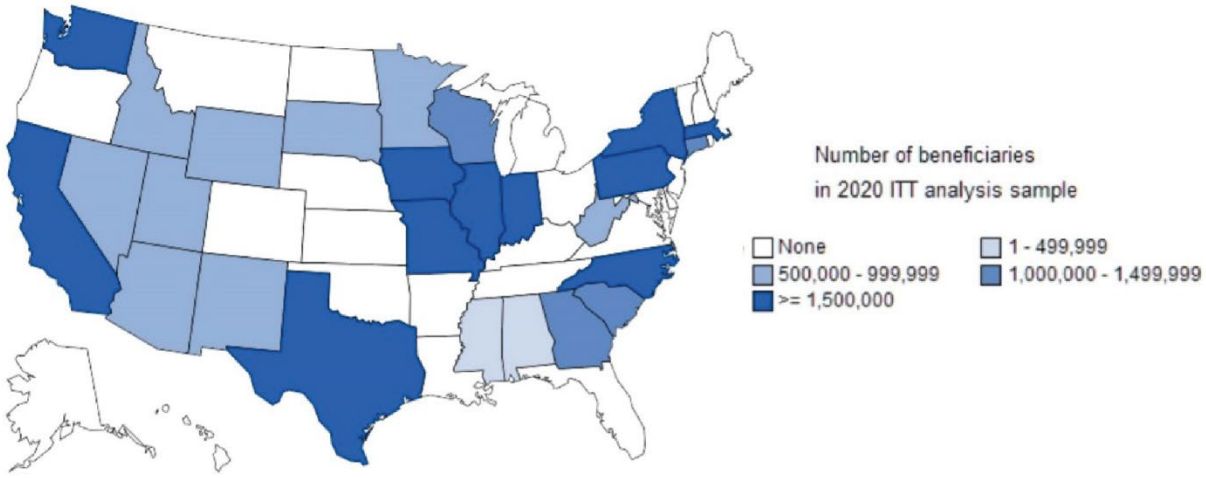
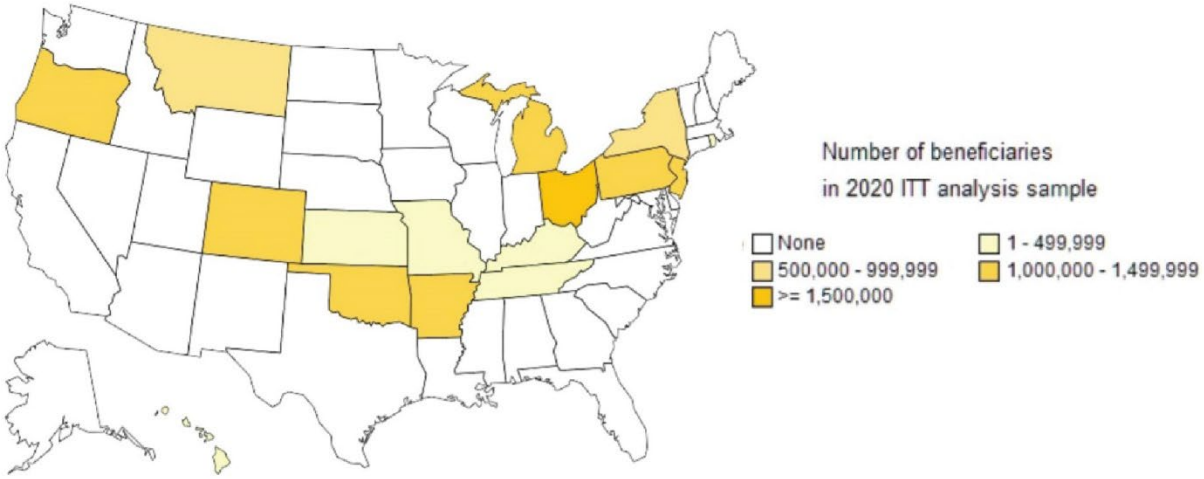
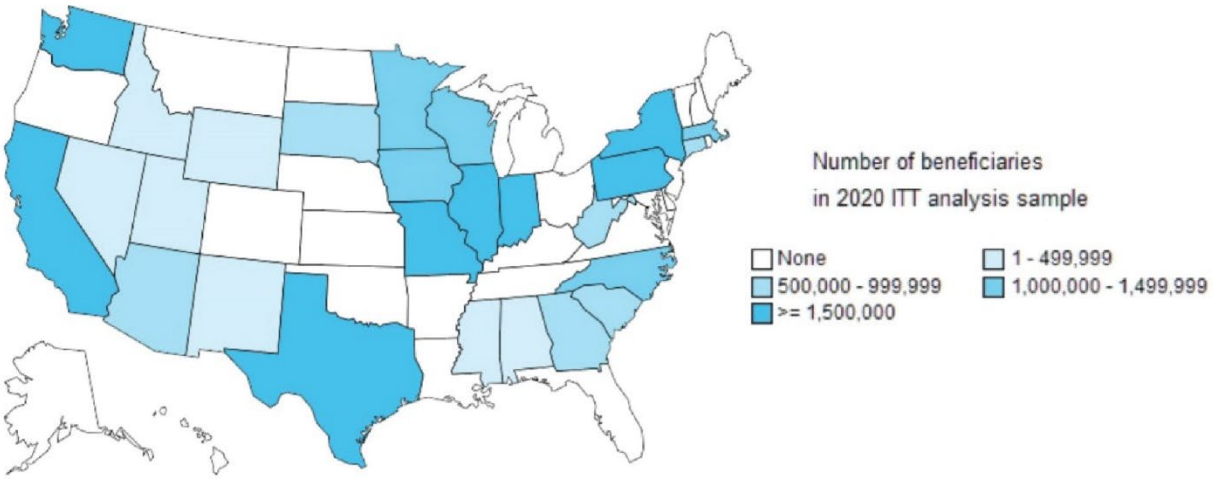


Figure 5.D.i. ***(continued)***

Panel C. Track 2 CPC+



Panel D. Track 2 comparison



Source: Mathematica's analysis of data on the number of Medicare beneficiaries assigned to CPC+ and comparison practices in 2020 from Medicare Enrollment Database. The state locations of practices are from IQVIA data. Eligible beneficiary months were weighted by the practice matching weight for comparison practices.

Note: CPC+ and comparison practices in Missouri, New York, and Pennsylvania are in different regions within the state. CPC+ regions include the Kansas City region in Kansas and Missouri, the North Hudson-Capital region in New York, and the Philadelphia region of Pennsylvania.

[a] Regions here are defined as states or contiguous counties.

CPC+ = Comprehensive Primary Care Plus; ITT = intent to treat.

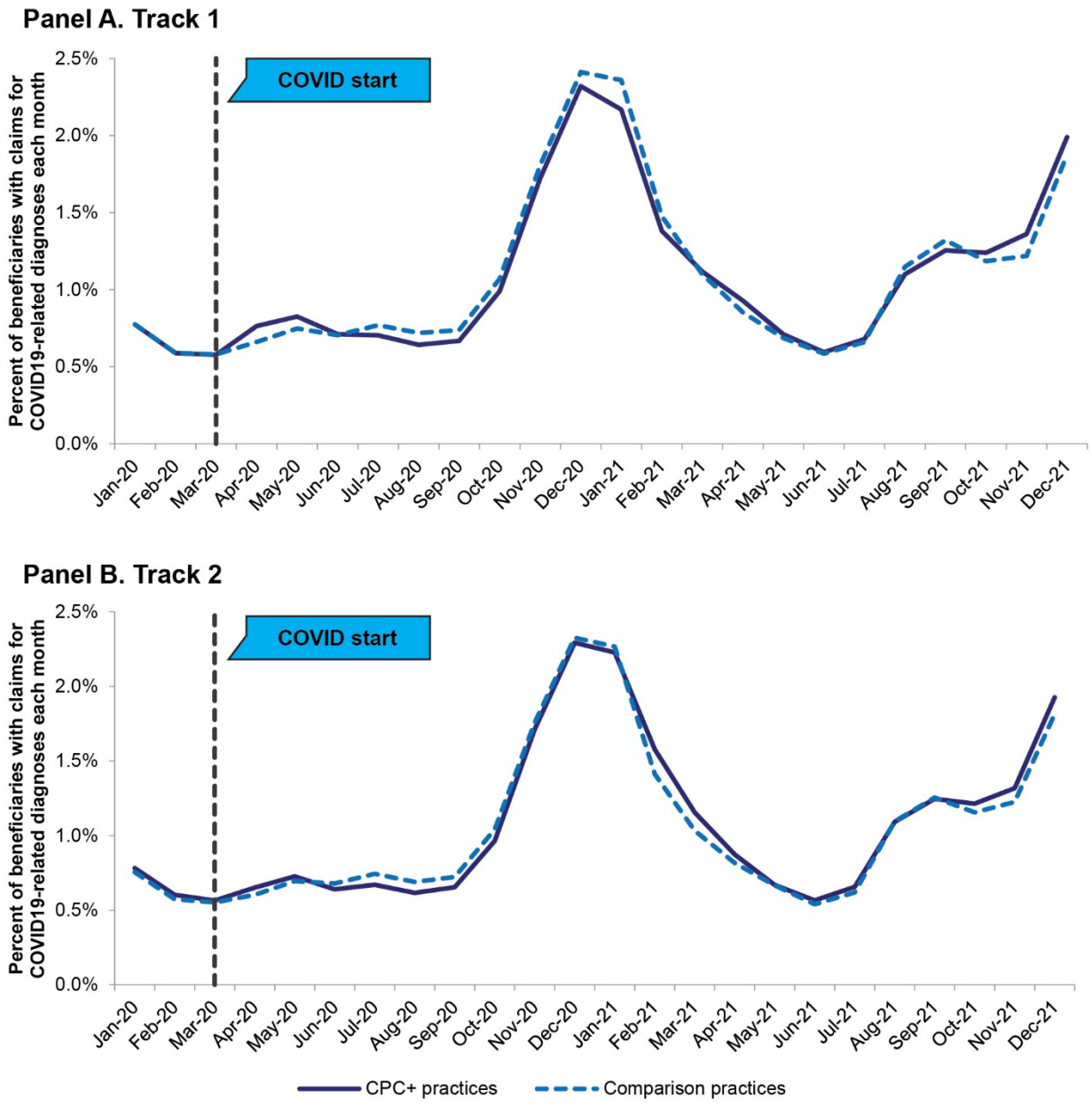
Supplement 2. Additional results for COVID-19-related diagnoses

A. Trends in COVID-19-related diagnoses

Figure 5.D.ii shows trends in COVID-19-related diagnoses for CPC+ and comparison practices from the beginning of 2020 through the end of 2021. Before March 2020, less than 1 percent of beneficiaries had a respiratory or coronavirus diagnosis each month, with the highest rates in January, a common time for respiratory viruses. Diagnoses were modestly higher in CPC+ practices in April 2020 (0.8 versus 0.7 percent in CPC+ versus comparison practices for Track 1, and 0.7 versus 0.6 percent for Track 2), at the beginning of the COVID-19 pandemic. This difference was somewhat more pronounced for Track 1 practices, likely driven by New Jersey, which was hit hard early in the pandemic and has a high proportion of Track 1 CPC+ practices. Starting in June and July 2020, diagnoses in comparison practices surpassed those in CPC+ practices and were higher through the rest of 2020. All practice groups reached their peak COVID-19-related diagnoses in late 2020 through early 2021, with a rate of about 2.5 percent of beneficiaries diagnosed per month, reflecting the wave around winter holidays in the first year of the pandemic. Diagnosis rates decreased afterwards for both practice groups through June 2021, reaching levels close to the beginning of the pandemic (about 1 percent each month), before rates increased again through the rest of 2021. By the end of 2021, amid the omicron surge, nearly 2 percent of CPC+ and comparison beneficiaries had a respiratory or coronavirus diagnosis each month in each track.

We observed similar trends in COVID-19-related diagnoses when we stratified CPC+ and comparison practices by SSP status, for both tracks (data not shown).

Figure 5.D.ii. Average rate of COVID-19-related diagnoses was similar among CPC+ and comparison beneficiaries across 2020 and 2021, but small differences in the timing of diagnoses reflected the geographic spread of COVID-19 across the country



Source: Medicare Part A and B claims data from 2020 and 2021.

Note: COVID-19-related diagnoses include COVID-19 diagnoses and respiratory conditions related to COVID-19 including viral pneumonia, acute bronchitis, lower respiratory infection, acute respiratory distress syndrome, and pneumonia because of COVID-19. See Bohl and Roozeboom-Baker (2020) for details. For comparison practices, percentages were weighted by matching weights.

CPC+ = Comprehensive Primary Care Plus.

B. Characteristics of beneficiaries with COVID-19-related diagnoses

In the full ITT analysis sample of CPC+ and comparison practices, beneficiaries diagnosed with COVID-19-related conditions accounted for 7 percent of the sample in 2020 and 9 percent of the sample in 2021. Diagnosis rates were similar in CPC+ and comparison groups and by track. All characteristics we examined were similar for CPC+ and comparison beneficiaries diagnosed with COVID-19-related conditions in each track (Table 5.D.i). Compared to the full ITT analysis sample, beneficiaries diagnosed with COVID-19 tended to have more chronic conditions, including diabetes (21–22 percent versus 15–16 percent), chronic obstructive pulmonary disease (18–19 percent versus 11–12 percent), congestive heart failure (19 percent versus 12 percent), and cardiovascular disease—including ischemic heart disease, acute myocardial infarction, and angina—(8–10 percent versus 5–6 percent), and higher hierarchical condition category (HCC) scores (1.7 versus 1.2).

Table 5.D.i. Characteristics of beneficiaries diagnosed with a COVID-19-related condition in 2020 and 2021 were similar for CPC+ and comparison practices in each track (characteristics shown as percentage, unless otherwise noted)

	Track 1				Track 2			
Measure	CPC+[a] (N = 160,361)	Comparison[a] (N = 544,248)	Difference	Standardized difference	CPC+[a] (N = 194,808)	Comparison[a] (N = 451,671)	Difference	Standardized difference
Age (mean)	72.9	73.0	-0.1	-0.01	72.9	72.8	0.1	0.01
Race								
White	88.5	87.9	0.6	0.02	88.2	87.7	0.5	0.02
Black	5.2	5.8	-0.6	-0.03	6.1	6.3	-0.2	-0.01
Other	6.3	6.3	0.0	0.00	5.7	6.1	-0.3	-0.01
Male	41.5	41.9	-0.4	-0.01	41.5	41.7	-0.2	-0.01
Original reason for Medicare eligibility								
Age	77.0	76.9	0.2	0.00	77.3	76.5	0.8	0.02
Disabled	22.0	22.0	0.0	0.00	21.8	22.4	-0.7	-0.02
ESRD	1.0	1.1	-0.2	-0.02	1.0	1.1	-0.2	-0.02
Chronic conditions								
Diabetes with chronic complications	21.3	21.6	-0.3	-0.01	21.8	21.8	0.0	0.00
Chronic obstructive pulmonary disease	19.2	18.5	0.7	0.02	18.6	18.2	0.4	0.01
Congestive heart failure	19.1	19.0	0.0	0.00	19.2	19.1	0.0	0.00
Ischemic heart disease, acute myocardial infarction, angina	9.5	8.6	1.0	0.04	9.2	8.7	0.5	0.02
HCC score[b]	1.7	1.7	0.0	-0.01	1.7	1.8	0.0	-0.02

Source: Mathematica's analysis of Medicare claims and enrollment data for January 2020 through December 2021.

Note: COVID-19-related diagnoses include COVID-19 diagnoses and respiratory conditions related to COVID-19 including viral pneumonia, acute bronchitis, lower respiratory infection, acute respiratory distress syndrome, and pneumonia due to COVID-19. See Bohl and Roozeboom-Baker (2020) for details. Characteristics were measured as of January 1, 2020, and January 1, 2021, using a two-year lookback period for chronic conditions.

[a] Means were weighted to account for (1) the share of the year for which the beneficiary's data were observed and (2) the matching (for beneficiaries in comparison practices only).

[b] HCC scores are a measure of risk for subsequent expenditures. The Centers for Medicare & Medicaid Services calculates them such that the average for the Medicare fee-for-service population nationally is 1.0. A patient with a risk score of 1.30 is predicted to have expenditures that would be approximately 30 percent above the average, whereas a patient with a risk score of 0.70 is expected to have expenditures that would be approximately 30 percent below the average.

CPC+ = Comprehensive Primary Care Plus; ESRD = end-stage renal disease; HCC = hierarchical condition category.

C. Direct effects on COVID-19-related diagnoses by SSP status

For both SSP and non-SSP practices, monthly rates of COVID-19-related diagnoses and rates of outpatient emergency department (ED) visits for COVID-19-related diagnoses were, on average, higher in 2021 than 2020 (Tables 5.D.ii and 5.D.iii). By contrast, utilization and expenditures for acute hospitalizations were somewhat lower in 2021 than in 2020. For SSP and non-SSP practices, the regional differences in the rates of COVID-19-related diagnoses and utilization and expenditures for COVID-19-related diagnoses were small. For example, the regional differences in acute hospitalizations for COVID-19-related diagnoses were less than 1 percent of the rate of all-cause acute hospitalizations for both SSP and non-SSP practices.

Table 5.D.ii. Within SSP and non-SSP practices, COVID-19-related diagnoses and outpatient ED visits for COVID-19-related diagnoses were higher in 2021 than in 2020, but hospitalizations were somewhat lower in 2021. Differences between practices in CPC+ and comparison regions were small (Track 1)

	Track 1 – SSP, Unadjusted means				Track 1 – SSP, Differences (SE)		Track 1 – Non-SSP, Unadjusted means				Track 1 – Non-SSP, Differences (SE)	
	CPC+	Comparison	Non-CPC+	Non-comparison	CPC+ vs. comparison	Non-CPC+ vs. non-comparison	CPC+	Comparison	Non-CPC+	Non-comparison	CPC+ vs. comparison	Non-CPC+ vs. non-comparison
Beneficiary claims with COVID-19-related diagnosis (percentage of beneficiaries with a claim each month)												
2020 (Mar–Dec)	1.1%	1.1%	1.3%	1.2%	0.00 p.p. (0.04 p.p.)	0.01 p.p. (0.04 p.p.)	0.9%	1.0%	1.1%	1.2%	-0.05 p.p. (0.03 p.p.)	-0.05 p.p.** (0.03 p.p.)
2021 (Jan–Dec)	1.3%	1.2%	1.4%	1.3%	0.03 p.p. (0.04 p.p.)	0.08 p.p.** (0.03 p.p.)	1.2%	1.2%	1.3%	1.4%	-0.03 p.p. (0.03 p.p.)	-0.05 p.p.** (0.02 p.p.)
2020–2021[a]	1.2%	1.2%	1.3%	1.3%	0.02 p.p. (0.04 p.p.)	0.05 p.p. (0.03 p.p.)	1.0%	1.1%	1.2%	1.3%	-0.04 p.p. (0.03 p.p.)	-0.05 p.p.** (0.02 p.p.)
Outpatient ED visits, including observation stays, with COVID-19-related diagnosis (per 1,000 beneficiaries per year)												
2020 (Mar–Dec)	9	11	11	13	-1.6*** (0.3)	-1.4*** (0.5)	11	13	13	15	-1.5*** (0.5)	-1.5*** (0.4)
2021 (Jan–Dec)	16	15	18	17	1.4*** (0.5)	0.5 (0.5)	16	17	19	19	-0.8 (0.6)	-0.4 (0.4)
2020–2021[a]	13	13	15	15	0.0 (0.4)	-0.4 (0.4)	14	15	16	17	-1.1** (0.5)	-0.9*** (0.4)
Acute hospitalizations with COVID-19-related diagnosis (per 1,000 beneficiaries per year)												
2020 (Mar–Dec)	19	21	23	24	-1.2** (0.6)	-0.7 (0.7)	18	19	22	23	-1.1* (0.7)	-0.6 (0.6)
2021 (Jan–Dec)	20	19	22	20	0.5 (0.4)	2.2*** (0.5)	18	18	21	21	-0.2 (0.5)	-0.5 (0.4)
2020–2021[a]	20	20	23	22	-0.2 (0.4)	0.9* (0.5)	18	19	22	22	-0.6 (0.5)	-0.5 (0.4)
Medicare inpatient expenditures for COVID-19-related diagnosis (per beneficiary per month)												
2020 (Mar–Dec)	$35	$38	$43	$46	-$3.6*** ($1.2)	-$3.1** ($1.6)	$31	$35	$39	$43	-$4.0*** ($1.4)	-$3.3** ($1.4)
2021 (Jan–Dec)	$35	$36	$42	$38	-$0.4 ($1.0)	$3.5*** ($1.2)	$32	$35	$37	$41	-$2.8** ($1.1)	-$3.6*** ($0.9)
2020–2021[a]	$35	$37	$43	$42	-$1.9** ($0.9)	$0.5 ($1.1)	$32	$35	$38	$42	-$3.4*** ($1.1)	-$3.5*** ($0.9)

Table 5.D.ii. ***(continued)***

	Track 1 – SSP, Unadjusted means				Track 1 – SSP, Differences (SE)		Track 1 – Non-SSP, Unadjusted means				Track 1 – Non-SSP, Differences (SE)	
	CPC+	Comparison	Non-CPC+	Non-comparison	CPC+ vs. comparison	Non-CPC+ vs. non-comparison	CPC+	Comparison	Non-CPC+	Non-comparison	CPC+ vs. comparison	Non-CPC+ vs. non-comparison
Unweighted sample sizes												
Number of practices	738	2,979	2,486	5,151			635	2,263	5,849	15,503		
Average number of beneficiaries per month	518,992	2,100,625	791,488	2,143,232			508,998	1,489,272	1,440,925	4,550,934		

Source: Mathematica's analysis of Medicare claims data from March 2020 through December 2021.

Note: COVID-19-related diagnoses include COVID-19 diagnoses and respiratory conditions related to COVID-19 including viral pneumonia, acute bronchitis, lower respiratory infection, acute respiratory distress syndrome, and pneumonia due to COVID-19. See Bohl and Roozeboom-Baker (2020) for details. Differences in the table are from time-series models run at the practice-month-year level that did not adjust for beneficiary or practice characteristics. For CPC+ practices, observations were weighted by the number of Medicare fee-for-service beneficiaries assigned to the practice during the month and year. For comparison practices, the weight is a product of the number of assigned beneficiaries and the matching weight. For non-CPC+ and non-comparison practices, we used a concentration weight constructed at the state-HRR level such that non-CPC+ practices had the same level of representation (in terms of beneficiary months) as CPC+ practices in the same state and HRR and SSP group, and non-comparison practices had the same level of representation as comparison practices in the same state and HRR and SSP group. Among non-CPC+ and non-comparison practices, we winsorized the weights at the 99th percentile.

[a] The 2020–2021 estimates used data from March 2020 to December 2021.

*/**/*** Significantly different from zero at the 0.10/0.05/0.01 levels, two-tailed test.

CPC+ = Comprehensive Primary Care Plus; ED = emergency department; HRR = hospital referral region; p.p. = percentage points; SE = standard error; SSP = Medicare Shared Savings Program.

Table 5.D.iii. Within SSP and non-SSP practices, COVID-19-related diagnoses and outpatient ED visits for COVID-19-related diagnoses were higher in 2021 than in 2020, but hospitalizations were somewhat lower in 2021. Differences between practices in CPC+ and comparison regions were small (Track 2)

	Track 2 – SSP, Unadjusted means				Track 2 – SSP, Differences (SE)		Track 2 – Non-SSP, Unadjusted means				Track 2 – Non-SSP, Differences (SE)	
	CPC+	Comparison	Non-CPC+	Non-comparison	CPC+	Comparison	Non-CPC+	Non-comparison	CPC+	Comparison	Non-CPC+	Non-comparison
Beneficiary claims with COVID-19-related diagnosis (percentage of beneficiaries with a claim each month)												
2020 (Mar–Dec)	1.0%	1.0%	1.1%	1.3%	-0.05 p.p. (0.03 p.p.)	-0.14 p.p.*** (0.03 p.p.)	0.9%	0.9%	1.1%	1.1%	-0.02 p.p. (0.03 p.p.)	-0.03 p.p. (0.02 p.p.)
2021 (Jan–Dec)	1.3%	1.2%	1.4%	1.3%	0.10 p.p.* (0.05 p.p.)	0.03 p.p. (0.03 p.p.)	1.2%	1.1%	1.3%	1.3%	0.02 p.p. (0.03 p.p.)	-0.01 p.p. (0.02 p.p.)
2020–2021[a]	1.1%	1.1%	1.3%	1.3%	0.03 p.p. (0.04 p.p.)	-0.05 p.p. (0.03 p.p.)	1.1%	1.1%	1.2%	1.2%	0.00 p.p. (0.03 p.p.)	-0.02 p.p. (0.02 p.p.)
Outpatient ED visits, including observation stays, with COVID-19-related diagnosis (per 1,000 beneficiaries per year)												
2020 (Mar–Dec)	10	11	12	13	-1.1 (0.7)	-1.7*** (0.6)	10	12	13	14	-1.1** (0.5)	-1.2*** (0.4)
2021 (Jan–Dec)	16	15	17	18	0.7 (0.6)	-0.2 (0.6)	16	15	20	19	0.2 (0.5)	0.7 (0.4)
2020–2021[a]	13	13	15	16	-0.1 (0.6)	-0.9* (0.5)	13	14	16	17	-0.4 (0.4)	-0.2 (0.3)
Acute hospitalizations with COVID-19-related diagnosis (per 1,000 beneficiaries per year)												
2020 (Mar–Dec)	20	21	24	25	-1.2 (0.9)	-1.7** (0.8)	17	19	21	22	-1.2* (0.7)	-1.1** (0.5)
2021 (Jan–Dec)	20	19	23	20	1.0* (0.5)	2.6*** (0.6)	18	18	20	21	-0.4 (0.5)	-0.6* (0.4)
2020–2021[a]	20	20	23	23	0.0 (0.6)	0.6 (0.6)	17	18	21	21	-0.7 (0.5)	-0.9** (0.4)
Medicare inpatient expenditures for COVID-19- related diagnosis (per beneficiary per month)												
2020 (Mar–Dec)	$34	$39	$42	$48	-$4.7*** ($1.8)	-$6.6*** ($1.7)	$31	$34	$38	$41	-$2.9** ($1.3)	-$2.7** ($1.1)
2021 (Jan–Dec)	$35	$35	$42	$39	$0.3 ($1.1)	$3.4** ($1.3)	$32	$33	$38	$40	-$1.7* ($1.0)	-$2.2*** ($0.9)
2020–2021[a]	$35	$37	$42	$43	-$1.9 ($1.2)	-$1.2 ($1.2)	$31	$34	$38	$40	-$2.2** ($1.0)	-$2.4*** ($0.8)

Table 5.D.iii. ***(continued)***

	Track 2 – SSP, Unadjusted means				Track 2 – SSP, Differences (SE)		Track 2 – Non-SSP, Unadjusted means				Track 2 – Non-SSP, Differences (SE)	
	CPC+	Comparison	Non-CPC+	Non-comparison	CPC+	Comparison	Non-CPC+	Non-comparison	CPC+	Comparison	Non-CPC+	Non-comparison
Unweighted sample sizes												
Number of practices	636	1,817	2,421	5,010			879	1,966	4,853	15,103		
Average number of beneficiaries per month	560,520	1,515,534	739,578	2,070,361			703,217	1,516,570	1,087,375	4,448,229		

Source: Mathematica's analysis of Medicare claims data from March 2020 through December 2021.

Note: COVID-19-related diagnoses include COVID-19 diagnoses and respiratory conditions related to COVID-19 including viral pneumonia, acute bronchitis, lower respiratory infection, acute respiratory distress syndrome, and pneumonia due to COVID-19. See Bohl and Roozeboom-Baker (2020) for details. Differences in the table are from time-series models run at the practice-month-year level that did not adjust for beneficiary or practice characteristics. For CPC+ practices, observations were weighted by the number of Medicare fee-for-service beneficiaries assigned to the practice during the month and year. For comparison practices, the weight is a product of the number of assigned beneficiaries and the matching weight. For non-CPC+ and non-comparison practices, we used a concentration weight constructed at the state-HRR level such that non-CPC+ practices had the same level of representation (in terms of beneficiary months) as CPC+ practices in the same state and HRR and SSP group, and non-comparison practices had the same level of representation as comparison practices in the same state and HRR and SSP group. Among non-CPC+ and non-comparison practices, we winsorized the weights at the 99th percentile.

[a] The 2020–2021 estimates used data from March 2020 to December 2021.

*/**/*** Significantly different from zero at the 0.10/0.05/0.01 levels, two-tailed test.

CPC+ = Comprehensive Primary Care Plus; ED = emergency department; HRR = hospital referral region; p.p. = percentage points; SE = standard error; SSP = Medicare Shared Savings Program.

Supplement 3. Detailed methods and additional results for the analysis of excess deaths

A. Study population, unit of observation, outcomes, and controls

Study population. Calculations of excess deaths in the Medicare FFS population included 89 state and hospital referral region (HRR) combinations (referred to as state-HRRs) containing CPC+ practices and 253 state-HRRs containing comparison practices, across both Tracks 1 and 2.[52] On average across the six years of the analysis (2016–2021), 7 million Medicare FFS beneficiaries were included in the analysis from CPC+ regions and 19 million Medicare FFS beneficiaries were included in the analysis from comparison regions.[53]

Unit of observation. The unit of observation in the regressions is a combination of state-HRR, month, year, CPC+ region, age group, race, and sex.[54] That is, each state-HRR has observations that correspond to the months from January 2016 to December 2021 (72 months), with each month further divided into combinations of age group, race, and sex (24 combinations), a total of 1,728 observations for each state-HRR (72 x 24).[55] Age groups included 0–64, 65–74, 75–84, and 85 and more. Race categories included White, Black, and all other (non-White or non-Black). Each observation was constructed using outcomes from beneficiaries who lived in that state-HRR during each month and were in the age, race, and sex category of the observation, according to Medicare enrollment data.

1. We applied weights to each observation with the goal of weighting the sample to represent the age, race, sex, and regional distribution of beneficiaries in the 2019 ITT analysis sample. We applied three weights to each observation, as follows:
2. All observations within a state-HRR received a weight that accounts for the share of beneficiaries contributed by that state-HRR to the total number of CPC+ or comparison beneficiaries in the ITT analysis sample for 2019, by track (for Track 1 and for Track 2 analyses) or by track and SSP status (for track- and SSP-specific analyses).56

All observations belonging to a particular age-race-sex combination received a weight that accounts for the share of beneficiaries contributed by that age-race-sex cell to the total number of CPC+ or comparison beneficiaries in the ITT analysis sample for 2019, by track (for Track 1 and for Track 2 analyses) or by track and SSP status (for track- and SSP-specific analyses).

[52] Most state-HRRs should contain only CPC+ or only comparison practices, and therefore have one observation per month in the data. However, there are seven state-HRR combinations from Missouri, New York, and Pennsylvania that include both CPC+ and comparison practices, and therefore have two observations, one with CPC+ practice-based weights and the other with comparison practice-based weights.

[53] Some of these Medicare FFS beneficiaries live in the seven state-HRR combinations that include both CPC+ and comparison practices and are counted in both the 7 million and 19 million beneficiaries in the CPC+ and comparison regions, respectively.

[54] The seven state-HRRs containing both CPC+ and comparison practices have duplicate observations—one with CPC+ practice-based weights and the other with comparison practice-based weights.

[55] The seven state-HRRs containing both CPC+ and comparison practices have a total of 2,880 observations (1,440 multiplied by 2).

[56] These weights are based on the state-HRR of the beneficiary's attributed practice for CPC+ or comparison beneficiaries in the ITT analysis sample for 2019, which may or may not differ from the state-HRR where the beneficiary resides.

All observations in a comparison group state-HRR received a weight based on the matching weights of comparison group practices in that state-HRR, by track (for Track 1 and for Track 2 analyses) or by track and SSP status (for track- and SSP-specific analyses).

The final weights for each observation in the regression are a product of these three weights. Results were very similar with and without weighting.

Outcomes. We analyzed the death rate in the state-HRR during each month using information on date of death from the Medicare enrollment data.

Control variables. We controlled for the age, race, and sex distribution of Medicare beneficiaries living in the state-HRR during the month.

B. Regression approach

In this section, we describe the model specification used to estimate excess deaths. In these regression models (one for each track and one for each track and SSP status combination), we adjusted for age, race, and sex to account for regional demographic differences. In Equation (5.D.1), let s index the state, h index the HRR, and c index whether the region (defined as the combination of state and HRR) contains CPC+ or comparison practices.[57] Further, let a index age, r race, and g sex of the Medicare FFS beneficiaries living in the state-HRR. Finally, let m index the month of year t, with t ranging from 2016 to 2021.

$$(5.D.1) \quad y_{\text{shcargmt}} = \gamma 1(\text{c=1}) + \kappa_{\text{m}} + \delta \text{t} + \sum_{x=1}^{24} \phi_x 1(t = 2020 \ or \ 2021, m = x) + 1(c = 1) \sum_{x=1}^{23} \tau_m 1(m = x) + \beta 1(c = 1)t + 1(c = 1) \sum_{x=1}^{24} \lambda_x 1(t = 2020 \ or \ 2021, m = x) + f(a, r, g) + \varepsilon_{\text{shcargmt}}$$

where

y_{shcargmt} represents death rates in state *s* and HRR *h* by treatment status *c* among Medicare FFS beneficiaries of age *a*, race *r*, sex *g* in month *m* of year *t*.

γ is the fixed effect for state-HRRs that contain CPC+ practices.

κ_m denotes fixed effects for calendar months that capture seasonality in deaths (with January as the reference month).

δ is the coefficient on a linear time trend *t* in years.

ϕ_x are fixed effects for each month *x* of 2020 and 2021 that capture any deviations from month-specific historical trends.

[57] The seven state-HRRs with both CPC+ and comparison practices have duplicate observations, one with CPC+ practice-based weights and the other with comparison practice-based weights.

The next set of terms interact month fixed effects, linear time trend, and fixed effects for each month of 2020 and 2021 with the fixed effect for state-HRRs that contain CPC+ practices.

τ_m captures the deviations in usual death rate seasonality across calendar months *m* for state-HRRs that contain CPC+ practices.

β allows for the linear time trend in years to be different for state-HRRs that contain CPC+ practices and state-HRRs that contain comparison practices.

λ_x allows for the deviations from the historical trend in each month *x* of 2020 and 2021 to also be different for state-HRRs that contain CPC+ practices and state-HRRs that contain comparison practices.

$f(a,r,g)$ consists of indicators for each age group (0–64, 65–74, 75–84, and 85 and older), race (White, Black, and all other), being male, interactions between each of these indicators with year *t*, interactions between each of these indicators with calendar months indicators (κ_m), interactions between each of these indicators with the indicator of state-HRRs that contain CPC+ practices $(1(c=1))$, and interactions between each of these indicators with each month of 2020 and 2021, where the last set of interactions allows the deviations from month-specific historical trends in 2020 and 2021 to be age-, race-, and sex-specific.

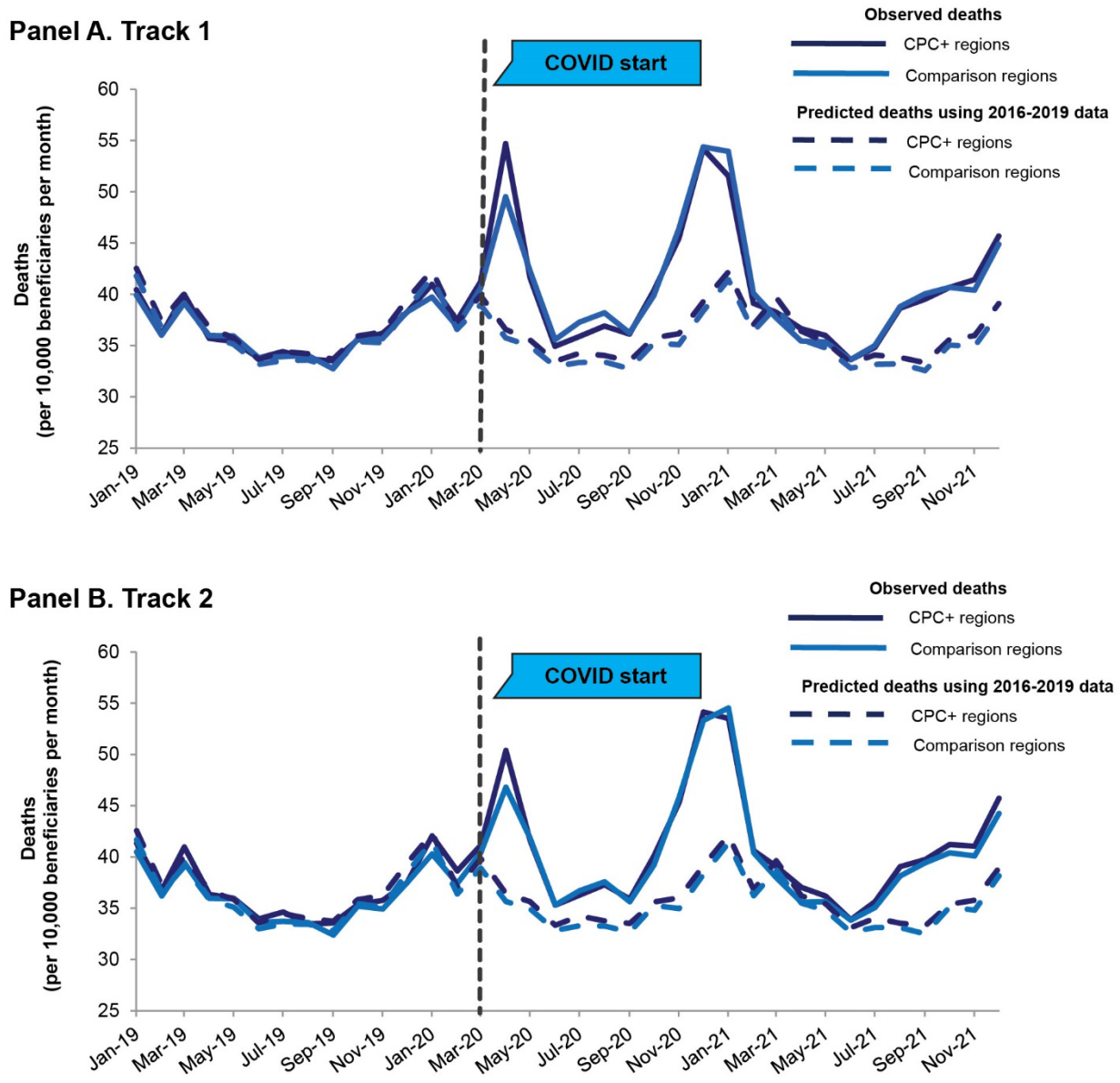
$\varepsilon_{shcargmt}$ is the idiosyncratic error term. It represents unexplained variability in the outcome variable for state *s* and HRR *h* by treatment status *c* among Medicare FFS beneficiaries of age *a*, race *r*, and sex *g* in month *m* of year *t*.

Standard errors are heteroskedasticity robust (Huber-White standard errors). In Equation (5.D.1), the coefficient λ_x captures the monthly difference in excess deaths between CPC+ and comparison regions. In the results presented in Table 5.D.2, we estimated a version of the model with a single λ capturing the average deviation in the historical trends for all months after the COVID-19 pandemic started (March 2020 through December 2021), and another version of the model, which allows λ to differ by each year of the pandemic (March through December 2020 and January through December 2021).

C. Trends in deaths in CPC+ and comparison regions

Predicted deaths using the regression model described in the previous section were very similar to observed deaths prior to 2020 and then diverged sharply after the start of the COVID-19 pandemic (Figure 5.D.iii, comparing dashed and solid lines). The differences in excess deaths between CPC+ and comparison regions fluctuated over time, peaking in April 2020 and becoming smaller afterwards. These differences between excess deaths in CPC+ and comparison regions were generally larger in regions weighted to represent Track 1 practices (Panel A) than in regions weighted to represent Track 2 practices (Panel B).

Figure 5.D.iii. CPC+ regions[a] had greater increases in deaths during some months, and smaller increases in deaths during other months than comparison regions, relative to predicted deaths if the COVID-19 pandemic had not occurred



Source: Mathematica's analysis of Medicare enrollment data from January 2016 through December 2021.

Note: Predicted deaths are based on models that are regression-adjusted for the distribution of age, race, and sex in the region. The models use data from 2016 through 2019 and project trends out through 2021 to predict deaths if the COVID-19 pandemic had not occurred. For calculating predicted deaths, we used observations at the state and HRR, month, year, age group, race, and sex levels. Each observation was weighted based on (1) the share of the 2019 ITT sample of CPC+ or comparison beneficiaries in that state-HRR, by track; (2) the share of the 2019 ITT sample of CPC+ or comparison beneficiaries in that age-race-sex cell, by track; and (3) the matching weights of comparison group practices in a state-HRR, by track.

[a] Regions are defined as the combination of state and HRR.

CPC+ = Comprehensive Primary Care Plus; HRR = hospital referral region; ITT = intent to treat.

D. Excess deaths in CPC+ and comparison regions, by track and by SSP status

Results when practices were weighted to represent SSP or non-SSP subgroups within each track were similar to results for the track overall. That is, CPC+ and comparison regions weighted to represent SSP subgroups in the regions had similar average excess deaths from March 2020 through December 2021, and the average excess deaths from March through December 2020 were somewhat higher than in 2021 in both CPC+ and comparison regions (Table 5.D.iv). Similarly, we obtained the same findings when weighting to represent non-SSP subgroups in the regions.

Table 5.D.iv. Excess deaths in 2020 and 2021 were similar among Medicare FFS beneficiaries in CPC+ and comparison regions, when weighted to represent SSP or non-SSP subgroups in Track 1 or Track 2[a]

Medicare FFS beneficiaries weighted to represent	Average excess deaths in March 2020–December 2021 in deaths per 10,000 beneficiaries per month (percent change from historical trends[a])							
	SSP, CPC+ regions[b]	SSP, Comparison regions[b]	SSP, Difference	SSP, 90% confidence interval	Non-SSP, CPC+ regions[b]	Non-SSP, Comparison regions[b]	Non-SSP, Difference	Non-SSP, 90% confidence interval
Track 1								
2020 (March–December)	7.7 (22%)	7.6 (22%)	0.1	(-0.8, 1.1)	6.3 (18%)	6.5 (19%)	-0.2	(-1.1, 0.6)
2021 (January–December)	4.2 (12%)	4.3 (12%)	-0.1	(-0.9, 0.7)	4.4 (12%)	4.9 (14%)	-0.4	(-1.0, 0.1)
2020–2021[c]	6.5 (18%)	6.4 (19%)	0.1	(-0.6, 0.7)	5.6 (16%)	5.9 (17%)	-0.3	(-0.9, 0.4)
Track 2								
2020 (March–December)	7.1 (20%)	7.4 (22%)	-0.3	(-1.1, 0.5)	6.2 (18%)	6.2 (18%)	0.0	(-1.2, 1.2)
2021 (January–December)	4.2 (12%)	4.2 (12%)	-0.1	(-0.8, 0.7)	4.5 (13%)	4.6 (13%)	0.0	(-0.8, 0.7)
2020–2021[c]	6.0 (17%)	6.2 (18%)	-0.2	(-0.8, 0.4)	5.6 (16%)	5.6 (16%)	0.0	(-1.0, 0.9)

Source: Mathematica's analysis of Medicare enrollment data from January 2016 through December 2021.

Note: Excess deaths are the difference between observed deaths in March 2020 through December 2021 and predicted deaths if COVID-19 had not occurred. Predicted deaths are based on models that are regression-adjusted for the distribution of age, race, and sex in the region. The models use data from 2016 through 2019 and project trends out through 2021 to predict deaths if the COVID-19 pandemic had not occurred. For calculating excess deaths, we used observations at the state and HRR, month, year, age group, race, and sex levels. Each observation was weighted based on (1) the share of the 2019 ITT sample of CPC+ or comparison beneficiaries in that state-HRR, by track and by SSP status; (2) the share of the 2019 ITT sample of CPC+ or comparison beneficiaries in that age-race-sex cell, by track and by SSP status; and (3) the matching weights of comparison group practices in a state-HRR, by track and by SSP status.

[a] To calculate these percentages, we divided the excess deaths in the region by the predicted deaths if COVID-19 had not occurred. Predicted deaths are based on regression models using data from 2016 through 2019.

[b] Regions defined as the combination of state and HRR.

[c] The 2020–2021 estimates used data from March 2020 to December 2021.

CPC+ = Comprehensive Primary Care Plus; FFS = fee for service; HRR = hospital referral region; ITT = intent to treat; SSP = Medicare Shared Savings Program.

Supplement 4. Creating control variables for regional excess mortality

This supplement provides detailed methods and additional results for creating a regional measure of excess deaths. As in Supplement 3, we estimated excess deaths as the difference between observed deaths in March through December 2021 and predicted deaths given historical trends from 2016 through 2019. Whereas the goal of the analyses presented in Supplement 3 was to compare excess deaths in CPC+ and comparison regions, the goal of the analyses presented in this supplement was to create a best estimate of excess deaths by region that can be used as a control variable in the impact analysis. Given this different objective, the methods presented in this supplement differ from Supplement 3 in two major ways:

1. The analyses presented here use Bayesian adjustment to "shrink" estimates of excess deaths toward the mean in a data-driven way. This provides more stable estimates of excess deaths. In regions with many cases, the number of predicted deaths will essentially equal the number of observed deaths in that region. In regions with few or no deaths, the predicted deaths will be closer to the average Medicare FFS death rate.
2. The regression models presented here do not use control variables because the output of the analyses presented here—the estimated excess deaths rates—will themselves become control variables in downstream difference-in-differences regression models that will adjust for beneficiary characteristics and practice fixed effects.

This supplement proceeds by (1) describing the study population, unit of observation, and outcome; (2) providing information about the regression approach used to calculate regional excess deaths; and (3) showing a comparison of predicted excess deaths relative to observed deaths.

A. Study population, unit of observation, and outcome

Study population. Calculations of excess deaths in the Medicare FFS population include 553 state-HRRs. Some contain either CPC+ or comparison practices, but some contain neither; these unrelated state-HRRs still provide useful information on national and state trends. On average across the six years of the analysis (2016–2021), 31 million Medicare FFS beneficiaries were included in the analysis. These 31 million Medicare FFS beneficiaries represent about 82 percent of the total Medicare FFS population.[58]

Unit of observation. The unit of observation in the regression models is the state-HRR-month-year. That is, each state-HRR has observations that correspond to the months from January 2016 to December 2021 (72 observations in total for each state-HRR). The state-HRR-month sample was constructed by taking unweighted counts of Medicare FFS beneficiaries who lived in that state-HRR during the month, according to Medicare enrollment data.

Outcome. We analyzed the death rate in the state-HRR during each month using date of death from the Medicare enrollment data.

[58] The Medicare FFS population included about 38 million beneficiaries each year from 2016 to 2020. https://www.kff.org/medicare/state-indicator/total-medicare-beneficiaries/?currentTimeframe=0&selectedRows=%7B%22wrapups%22:%7B%22united-states%22:%7B%7D%7D%7D&sortModel=%7B%22colId%22:%22Location%22,%22sort%22:%22asc%22%7D

B. Regression approach

In this section, we describe the model specification used to estimate the regional measures of excess deaths. We used a Bayesian hierarchical logistic regression model, which uses partial pooling to "shrink" smaller regions toward the mean in a data-driven way. In Equation (5.D.2), let *s* index the state, *h* index the HRR, and *m* index the month of year *t* with *t* ranging from 2016 to 2021.

$$
\begin{aligned}
(5.D.2) \qquad & y_{shmt} \sim \text{Binomial}\left(n_{shmt}, p_{shmt}\right) \\
& \text{logit}\left(p_{shmt}\right) = \alpha_0 + \alpha_s + \alpha_{sh} + \kappa_{0m} + \kappa_{sm} + \kappa_{shm} + \beta_{mt} + \delta_0 t + \delta_s t + \delta_{sh} t \\
& \qquad + \phi_{0m\text{covid}} + \phi_{sm\text{covid}} + \phi_{shm\text{covid}}
\end{aligned}
$$

where

y_{shmt} represents the number of deaths in state *s* and HRR *h* in month *m* of year *t*.

n_{shmt} is the number of Medicare FFS beneficiaries residing in state *s* and HRR *h* in month *m* of year *t*.

p_{shmt} is the death rate in state *s* and HRR *h* in month *m* of year *t*.

α_0 is an intercept term estimating the average national death rate.

α_s denotes a series of state-level random intercepts, allowing average death rate to vary by state.

α_{sh} denotes a series of state-HRR-level random intercepts, allowing average death rate to vary by state-HRR.

κ_{0m} denotes random effects for each of the 12 calendar months that capture national seasonality in deaths.

κ_{sm} denotes random effects that allow for state-level seasonality that differs from national seasonality.

κ_{shm} denotes random effects that allow for state-HRR-level seasonality that differs from state and national seasonality.

β_{mt} denotes random effects for each of the 72 months in the sample, capturing national monthly death rates that differ from a typical year's seasonality.

δ_0 is the coefficient on a national linear time trend *t* in years.

δ_s denotes a series of state-level random linear time slopes that allow a state's linear trend in death rates to differ from the national average.

δ_{sh} denotes a series of state-HRR-level random linear time slopes that allow a state-HRR's linear trend in death rates to differ from the average in the state.

$\phi_{0m\text{covid}}$ denotes a series of random effects for each month m during the COVID-19 pandemic from March 2020 to December 2021 that captures national deviations from month-specific historical trends during the pandemic.

$\phi_{sm\text{covid}}$ denotes a series of random effects for each state s and month m during the COVID-19 pandemic from March 2020 to December 2021 that captures state deviations from the month-specific national average during the pandemic.

$\phi_{shm\text{covid}}$ denotes a series of random effects for each state-HRR *sh* and month *m* during the COVID-19 pandemic from March 2020 to December 2021 that captures state-HRR deviations from the month-specific state average during the pandemic.

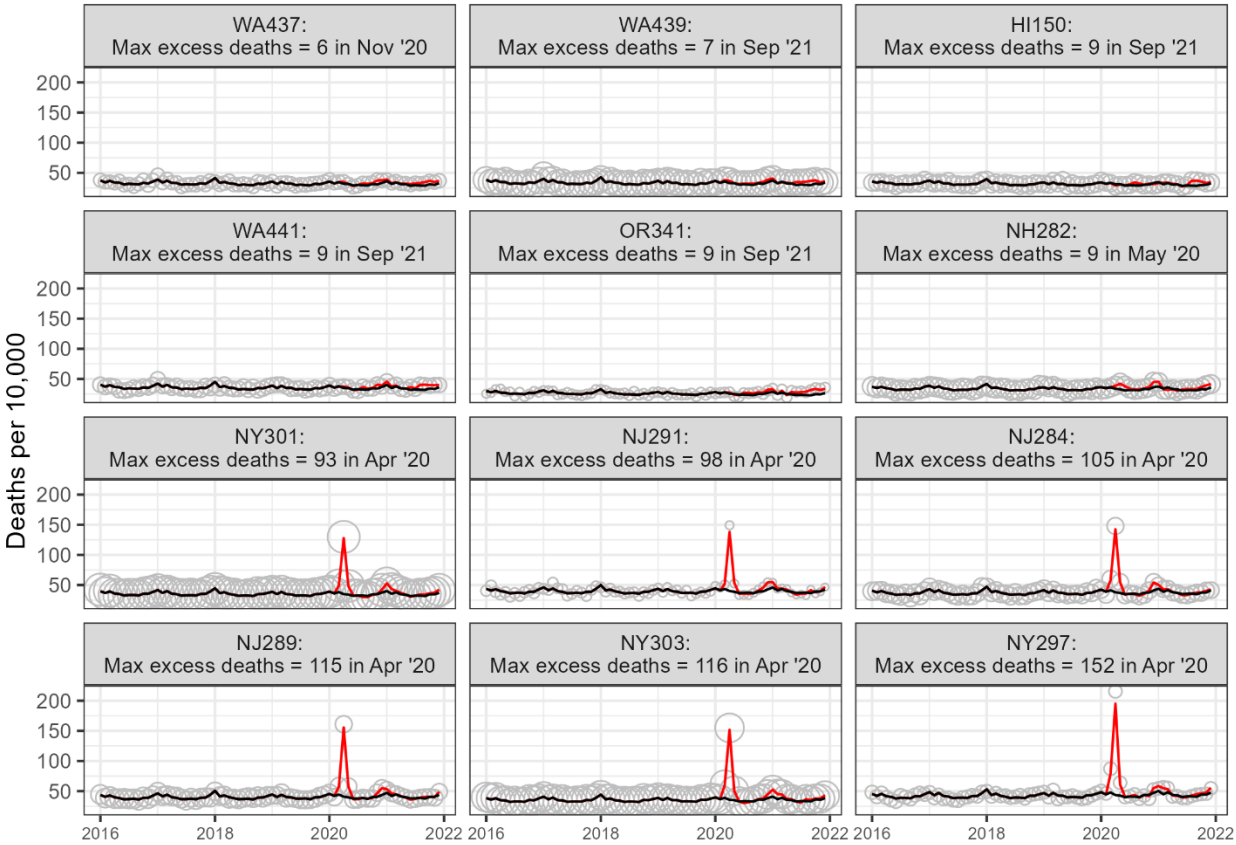
We estimated the excess deaths for each state *s* and HRR *h* in each month *x* as $\phi_{0x} + \phi_{sx} + \phi_{shx}$, that is, the sum of the national excess mortality rate for the month, plus the state excess mortality rate for the month, plus the state-HRR excess mortality rate for the month.

C. Estimates of excess deaths

Using Bayesian methods allowed us to smooth noisy estimates of deaths for small regions (for example, see the small state-HRR "OR341" in Figure 5.D.iv) while still very accurately capturing peaks in deaths due to COVID-19 (see how closely the peaks in red lines overlap with circles in Figure 5.D.iv).[59]

[59] The Bayesian model was fit with the probabilistic programming language Stan (Stan Development Team. "Stan Modeling Language Users Guide and Reference Manual, 2.27." 2019. https://mc-stan.org). Although 10 iterations out of 16,000 posterior draws had divergent transitions—an issue that could affect predictive performance—we think these computational challenges had a negligible impact on the estimates of excess death rates used in this Appendix.

Figure 5.D.iv. Predicted and observed death rates for 12 example regions with the lowest and highest peak excess deaths from 2020–2021, for which predicted deaths using Bayesian random effects models were very similar to observed deaths when sample sizes were large and appropriately different when sample sizes were small



Source: Mathematica's analysis of Medicare enrollment data from January 2016 through December 2021

Note: The red line represents predicted deaths during the COVID-19 pandemic using a random effects model; the black line represents predicted deaths if the COVID-19 pandemic had not occurred (using only data from 2016 through 2019 to predict deaths in 2020–2021) using the same random effects model; and the circles represent observed deaths, with larger circles denoting that more beneficiaries were included in the observed rates. The graph shows that the predicted estimates during the COVID-19 pandemic are very similar to those observed. For smaller regions like OR341, the random effects model "shrinks" the predicted estimates toward the mean in a data-driven way.

HI = Hawaii; NH = New Hampshire; NJ = New Jersey; NY = New York; OR = Oregon; WA = Washington.

Supplement 5. Detailed methods to account for any differences using COVID-19-related controls

A. Study population, unit of observation, outcomes, and controls

Study population. We applied the regression models to the 2017 CPC+ Starters and comparison regions. The sample includes Track 1 and Track 2 CPC+ and comparison practices as well as unselected practices in Track 1 and Track 2 CPC+ and comparison regions.

Unit of observation. The unit of observation in the regression models is the practice-month-year. That is, each practice has observations that correspond to the months from January 2019 to December 2021 (36 observations in total for each practice). The practice-month-year sample was constructed by aggregating beneficiary-month observations, including enrollment and outcomes of interest, to assigned practices. We aggregated to the practice level to reduce the time it takes to run the models. We included Medicare FFS beneficiaries in the ITT analysis sample who were eligible during a given month from 2019 to 2021 and assigned to one of the four practice types included in the analysis (CPC+, comparison, non-CPC+, and non-comparison). In the aggregation process, we weighted beneficiaries assigned to comparison practices by a matching weight, which ensures CPC+ and comparison practices were comparable in terms of baseline characteristics (as required for the CPC+ impact evaluation). For beneficiaries in non-CPC+ or non-comparison practices, beneficiaries were weighted by the concentration of CPC+ and comparison practices in the same state and HRR and SSP group prior to practice-level aggregation, which ensures non-CPC+ practices had the same level of representation (in terms of beneficiary months) as CPC+ practices in the same state and HRR and SSP group, and non-comparison practices had the same level of representation as comparison practices in the same state and HRR and SSP group.[60] To minimize the potential for results being driven by extreme weight values, we winsorized the weights at the 99th percentile.

Outcomes. We analyzed six key outcomes for Medicare FFS beneficiaries:

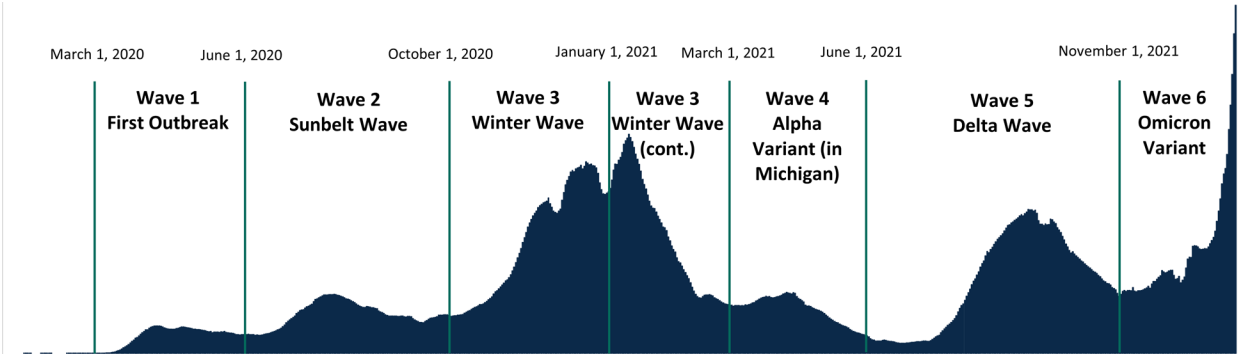
- Medicare expenditures without CMS's enhanced payments, in dollars per beneficiary per month
- Annualized number of acute hospitalizations per 1,000 beneficiaries
- Expenditures on acute hospitalizations, in dollars per beneficiary per month
- Annualized number of outpatient ED visits per 1,000 beneficiaries
- Expenditures on outpatient ED visits, in dollars per beneficiary per month
- Annualized number of urgent care center visits per 1,000 beneficiaries

Control variables. We included practice-level averages of beneficiary hierarchical condition category (HCC) scores, measured at the start of 2016 (the baseline period for the CPC+ impact evaluation) and, in the model section described in Section B of this supplement, the following regional COVID-19-related controls, based on practices' geographic location:

[60] The only exception to the balanced representation at the state-HRR level is for state-HRRs that had only CPC+ or comparison practices, in which case there is no representation of nonparticipating practices or unselected practices in those specific state-HRRs. We adjusted the concentration weight for practices that were in the same state for such cases so that the representation at state level is still balanced.

- Excess deaths in the state-HRR averaged during each "wave" of the pandemic in 2020 through 2021. We defined a total of seven waves: three waves in 2020 and four waves in 2021. Specifically, in 2020, the three waves were Wave 1 (March–May 2020), Wave 2 (June–September 2020), and Wave 3 (October–December 2020). In 2021, the four waves were Wave 3 (continued from 2020; January–February 2021), Wave 4 (March–May 2021), Wave 5 (June–October 2021), and Wave 6 (November–December 2021) (Figure 5.D.v).
- The maximum excess death estimate in the state-HRR in the year and the wave in which the maximum value occurred (for 2020 and 2021, separately)
- Pandemic Vulnerability Index, measured at the county level and averaged during each wave of the pandemic
- Government Response Index, measured at the state level using an average of the response across the year (for 2020 and 2021, separately)
- Social Vulnerability Index, measured at the census tract level using census information from 2018

Figure 5.D.v. Six waves of the pandemic and their relation to COVID-19-related cases from 2020 through 2021



Source: Mathematica's analysis of the Centers for Disease Control and Prevention's COVID-19 Case Surveillance Public Use Data with Geography.

Note: The solid blue represents daily COVID-19 cases in the United States from February 1, 2020, to December 31, 2021. The green vertical lines represent the start and end date for each wave definition used to specify some of the COVID-19 controls. The figure shows how daily COVID-19 cases in the United States correspond to each wave definition.

B. Regression approach for differences between unselected practices in CPC+ and comparison regions and differences between CPC+ and comparison practices

In this section, we describe the model specifications used to calculate differences between unselected practices in CPC+ and comparison regions and between CPC+ and comparison practices with and without COVID controls.

To estimate differences between unselected practices in CPC+ and comparison regions, we ran the models shown in Equation (5.D.3) (a yearly model) and Equation (5.D.4) (a cumulative model). Let j index the practice and t index time, where t ranges from 1 to 36, with values 1 to 12 denoting the months in 2019, values 13 to 24 denoting the months in 2020, and values 25 to 36 denoting the months in 2021. The models take the following form:

$$(5.D.3) \qquad y_{jt} = \alpha + \delta_t month_t + \gamma_t year_t + b_j + \theta_t z_j year_t + \beta_t^B X_j^B year_t + \beta_t^C X_j^C p_t + \varepsilon_{jt}$$

$$(5.D.4) \qquad y_{jt} = \alpha + \delta_t month_t + \gamma_t year_t + b_j + \theta z_j post2020 + \beta_t^B X_j^B year_t + \beta_t^C X_j^C p_t + \varepsilon_{jt}$$

where

y_{jt} represents a claims-based outcome variable in practice *j,* in period *t.*

$month_t$ is a set of calendar month indicators (February, March, …, December) excluding January as the reference category, which takes the value of 1 during the month that corresponds to the specific calendar month and 0 otherwise.

$year_t$ is a set of calendar year indicators (2020, 2021) excluding 2019 as the reference category, which takes the value of 1 during the months that correspond to each specific calendar year and 0 otherwise.

b_j is a practice-level fixed effect for practice *j*, which controls for all time-invariant practice characteristics.

z_j is a binary indicator of practice being in a CPC+ region; the indicator takes the value of 1 if practice *j* is an unselected practice in the CPC+ region and 0 otherwise. The main effect of this indicator is not identified in this equation because it is collinear with the practice fixed effects.

$post2020$ is an indicator that takes the value of 1 during the months in years 2020 and 2021 and 0 otherwise (applicable to the cumulative model in Equation 5.D.4).

X_j^B is a vector of beneficiary characteristics—age, Chronic Conditions Warehouse Alzheimer's and dementia condition flag, and HCC scores—measured at the start of 2016 (the baseline period for the CPC+ impact evaluation) and aggregated to the practice level. We interacted each beneficiary characteristic control in X_j^B with year indicators for 2020 and 2021 (that is, $year_t$). The main effects of these average beneficiary characteristics are not identified because they are collinear with the practice fixed effects.

X_j^C is a vector of COVID-19-related regional controls (used in the models with COVID-19-related controls), which are outlined in Table 5.D.5. These regional controls comprise the following variables averaged during the months within each wave of the pandemic in 2020 through 2021: excess deaths in the state-HRR of practice *j,* measured separately for each of the seven pandemic waves in our study, and the Pandemic Vulnerability Index in the county of practice j, measured separately each wave. We also included the maximum monthly excess deaths in the state-HRR of practice *j,* separately for years 2020 and 2021; the wave that the maximum value occurred, separately for years

2020 and 2021; the 2020 and 2021 Government Response Index in the state of practice *j;* and the 2018 Social Vulnerability Index in the census tract of practice *j*. We interacted each of the COVID-19-related controls with 2020 and 2021 indicators, p_t . Specifically, we interacted the 2018 Social Vulnerability Index with year indicators for 2020 and 2021; we interacted all the other COVID-19-related controls with the contemporaneous year (2020 or 2021). The main effects of these COVID-19-related controls are not identified because they are collinear with the practice fixed effects.

ε_{jt} is the idiosyncratic error term. It represents the effect of the unobserved factors that can influence the outcome variable for practice *j,* during period *t.*

In the yearly model (Equation 5.D.3), θ_t captures the differences in outcomes separately for the months in 2020 and the months in 2021, relative to that difference in 2019, adjusting for beneficiary risk scores, practice fixed effects, and COVID-19-related controls (if applicable). In the cumulative model (Equation 5.D.4), the coefficient θ captures the average differences in outcomes for the months in 2020 and 2021 (together), relative to that difference in 2019.

Standard errors are heteroskedasticity robust (Huber-White standard errors) and clustered at the practice level. For comparison-group observations, we applied weights that are equal to the product of a practice's total FFS Medicare enrollment during the given month-year so that practices with a higher number of assigned beneficiaries receive relatively more weight than practices with fewer assigned beneficiaries in the same period, and the matching weight. For practices in the CPC+ group, we needed only the enrollment weight because the matching weight for each CPC+ practice is 1.

To estimate differences between CPC+ and comparison practices (as opposed to non-CPC+ and non-comparison practices), we ran the same form of model as in Equations (5.D.3) and (5.D.4). All model elements are defined the same, except that now z_j is a binary indicator of being a CPC+ practice—the indicator takes the value of 1 if practice j is a CPC+ practice and 0 otherwise. Again, the main effect of this indicator is not identified in this equation because it is collinear with the practice fixed effects.

Supplement 6. Regional distribution of COVID-19 vaccination rates

Figure 5.D.vi shows the regional distribution of COVID-19 vaccination rates among counties where Track 1 and Track 2 CPC+ and comparison practices are located, with darker colors indicating states with a higher percentage of the counties' population of people 65 and older (65+) who were fully vaccinated against COVID-19. In each panel of the figure, we show the fully vaccinated rates (that is, having completed a single-dose or a two-dose vaccine as described in Supplement 7) in the middle of 2021 (July 2021) on the left and by the end of 2021 (December 2021) on the right, respectively.

In July 2021, the states with the largest percentage of the 65+ population fully vaccinated against COVID-19 in CPC+ counties were New York, Kentucky, Oregon, Pennsylvania, and Kansas (Figure 5.D.vi). Among comparison practices, the states with the largest percentage of the 65+ population fully vaccinated against COVID-19 in a county were Washington, Connecticut, Minnesota, Wisconsin, and Iowa. By end of 2021, the pattern across states we saw in July remained; however, vaccination rates increased such that most states had at least an average of 80 percent of the 65+ population vaccinated against COVID-19 in their counties.

Figure 5.D.vi. From July 2021 to December 2021, regional variation in the percentage of 65+ population being fully vaccinated against COVID-19 decreased

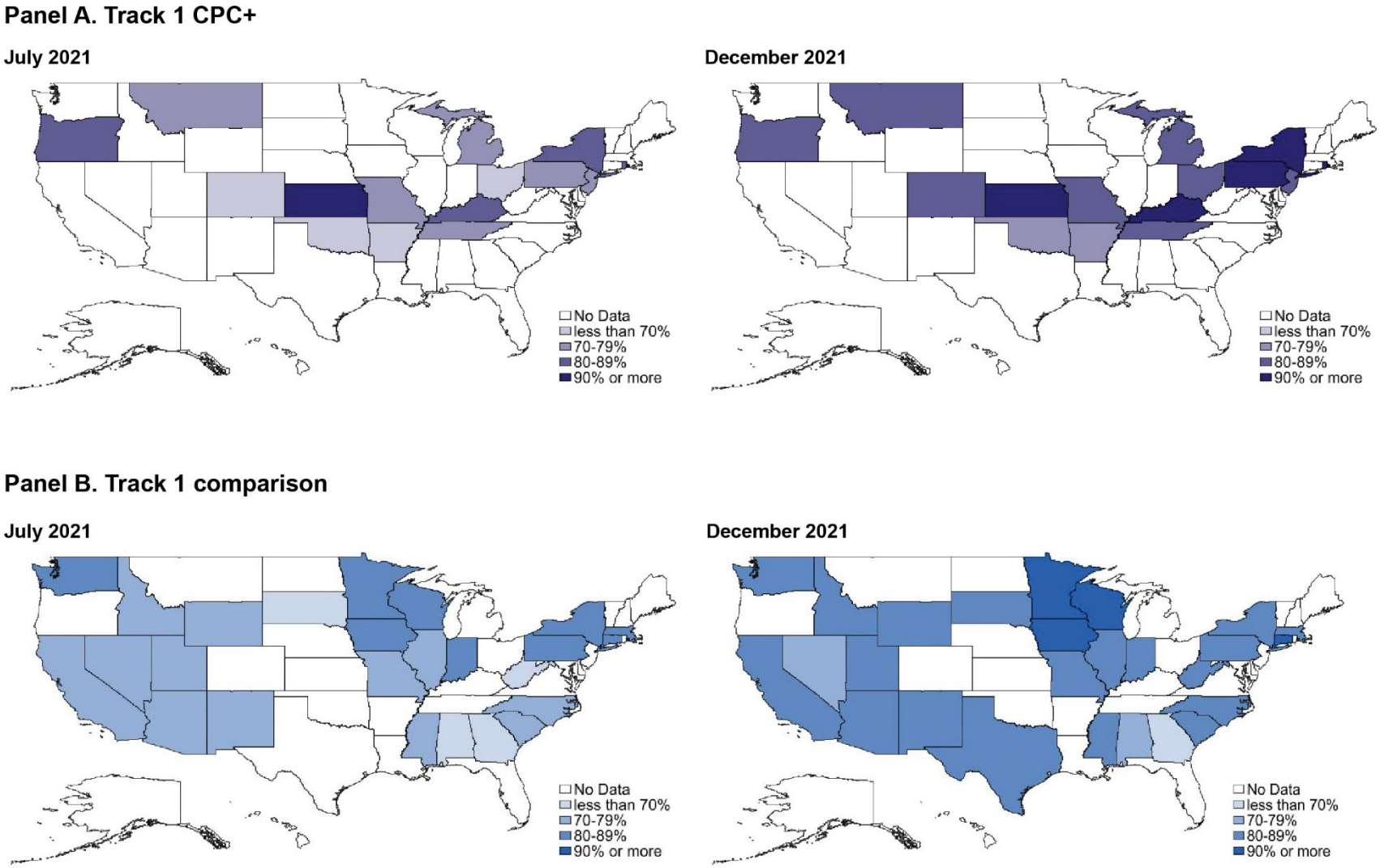
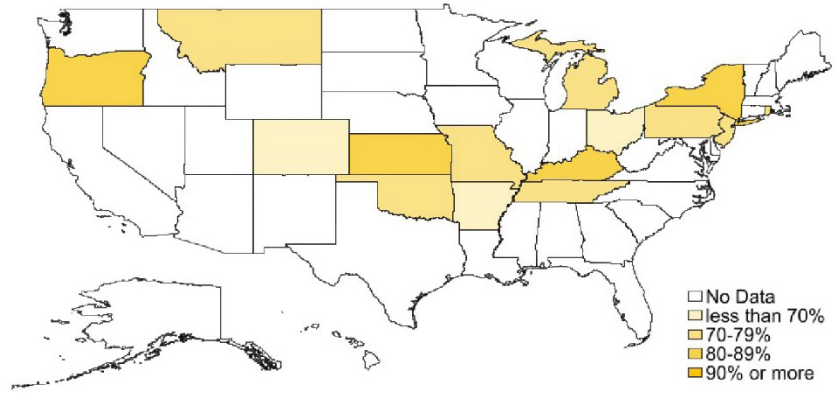


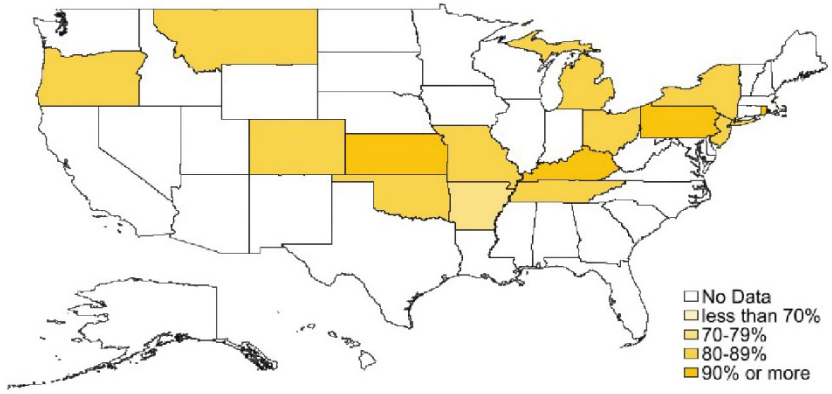
Figure 5.D.vi. ***(continued)***

Panel C. Track 2 CPC+

July 2021

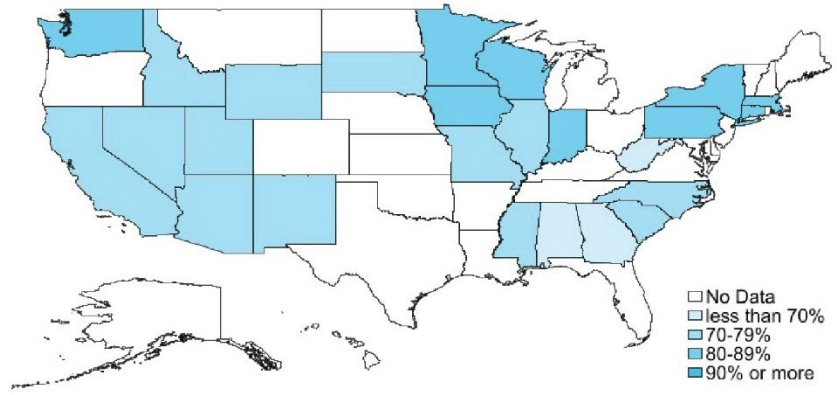


December 2021

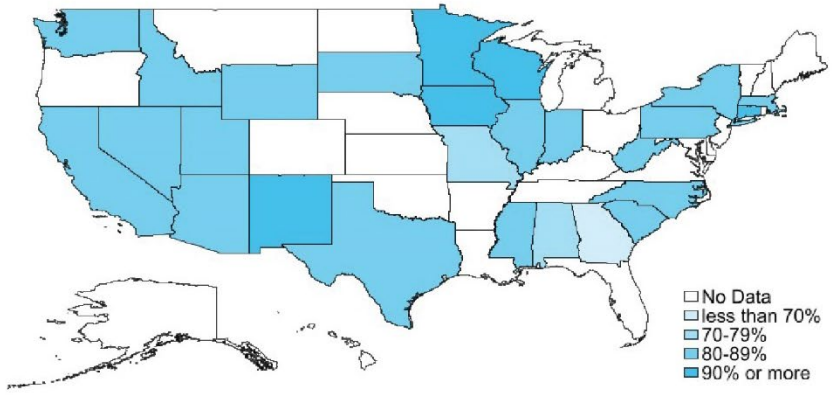


Panel D. Track 2 comparison

July 2021



December 2021



Source: Mathematica's analysis of the Centers for Disease Control and Prevention's county-level COVID-19 vaccination rates (July 2021 and December 2021) combined with data on the number of Medicare FFS beneficiaries assigned to selected and unselected practices in 2021 (from Medicare claims and Medicare Enrollment Database).

Figure 5.D.vi. ***(continued)***

Note: These figures show the distribution of COVID-19 vaccination rates across states, in counties of CPC+ and comparison practices as well as in the counties of non-CPC+ and non-comparison practices by track. We show the vaccination rates for the middle of 2021 (end of July 2021) and by the end of 2021 (end of December 2021). We use county-level COVID-19 vaccination rates defined as percentage of a county's 65+ population being fully vaccinated against COVID-19 (that is, per CDC, if they received one dose of a single-dose vaccine such as Johnson & Johnson, or two doses on of either an mRNA or protein-based series such as Moderna and Pfizer). Data on the July 2021 vaccination rates were missing for Hawaii and Texas; data on the December 2021 vaccination rates were missing for Hawaii. For CPC+ practices, observations were weighted by the number of Medicare FFS beneficiaries assigned to the practice during the month and year. For comparison practices, the weight is a product of the number of assigned beneficiaries and the matching weight.

CPC+ = Comprehensive Primary Care Plus; FFS = fee-for-service.

Supplement 7. Additional results on the total effects of COVID-19 for claims-based outcomes

In this section, we show additional results from our analysis for the total effects of COVID-19. Section A shows monthly trends in Medicare expenditures (the primary outcome in the impact analysis) and UCC visits (with a distinctly different pattern from other outcomes) for non-CPC+ and non-comparison practices from the beginning of 2019 through the end of 2021. Section B illustrates the year-to-year and two-year–to–two-year variation in the unadjusted regional differences of acute care utilization and expenditures among unselected practices from 2016 to 2021. Section C shows the adjusted regional differences among selected practices, with and without COVID-19-related controls. Section D describes sensitivity checks that include COVID-19 vaccination rates as additional regional COVID-19-related controls. Section E presents updated findings from using the triple-differences model with data from 2019 through 2021, another option to account for regional differences due to COVID-19.

A. Trends in Medicare expenditures and UCC visits from 2019 to 2021

Medicare expenditures dipped by 30 percent in early 2020, before returning to and reaching the pre-pandemic levels in late 2020 through the end of 2021 for both non-CPC+ and non-comparison practices (Figure 5.D.vii). Reductions in Medicare expenditures were on average larger for non-CPC+ practices than for non-comparison practices in both 2020 and 2021 (relative to 2019) but the differences were small, with larger differences in the first half of each year and smaller differences later in the year.

In the case of UCC visits, there was a similar but larger dip (by more than 50 percent) in early 2020 for both practice groups, but in contrast with Medicare expenditures, UCC visits increased thereafter and surpassed the pre-pandemic levels by the end of 2020 (Figure 5.D.viii). In another difference, growth in UCC visits was on average larger for non-CPC+ practices than non-comparison practices during the pandemic period (relative to 2019), particularly toward the end of 2021.

Figure 5.D.vii. Reductions in Medicare expenditures were larger for non-CPC+ practices than for non-comparison practices during the pandemic period, but regional differences in the reductions were small (relative to 2019)

Panel A. Track 1

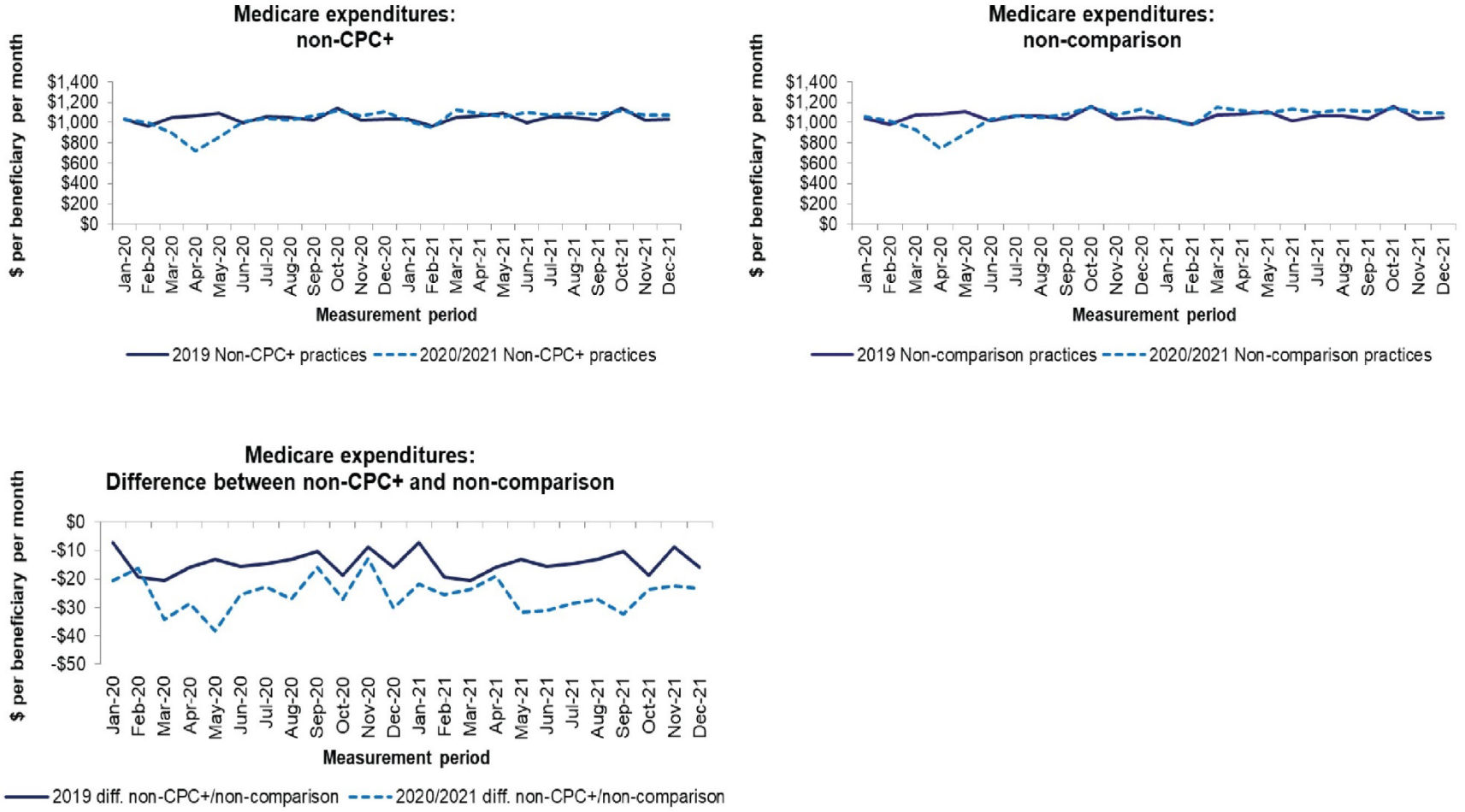
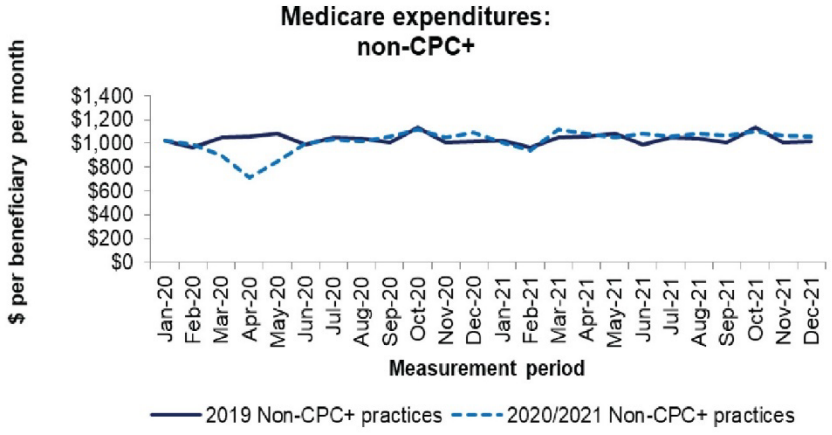
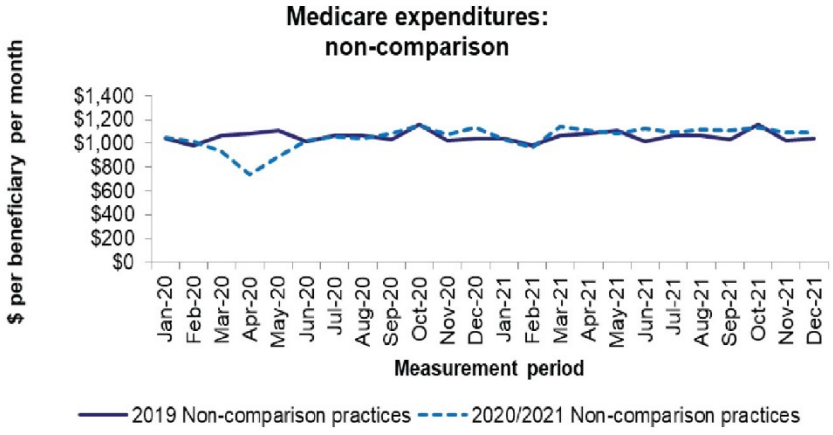
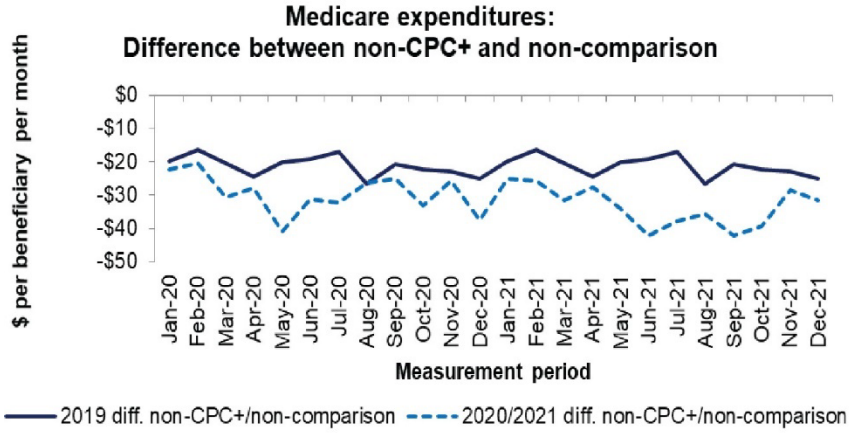


Figure 5.D.vii. ***(continued)***

Panel B. Track 2







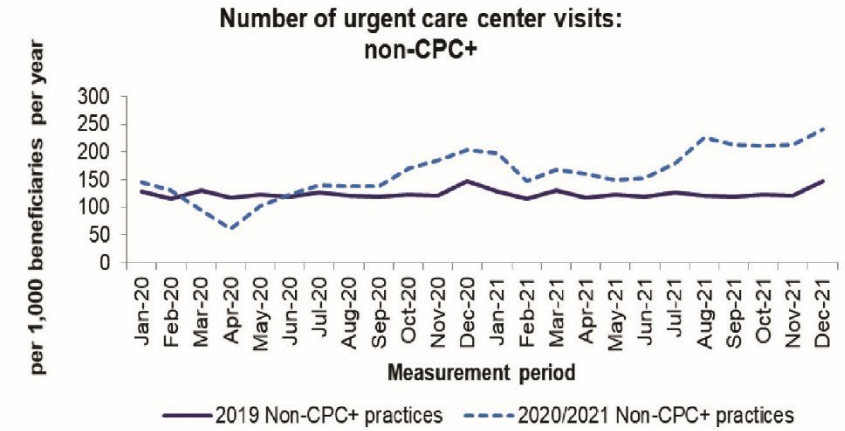
Source: Mathematica's analysis of Medicare Part A and B claims data from January 2019 through December 2021.

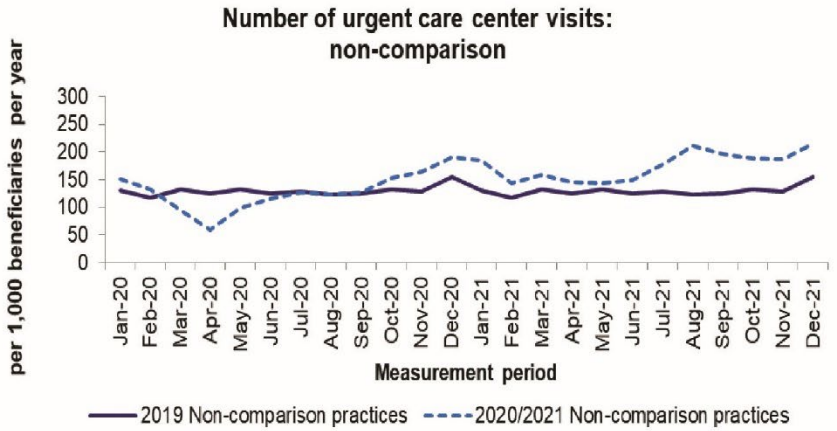
Note: Lines show Medicare expenditures without the Centers for Medicare & Medicaid Services' enhanced payments for non-CPC+ and non-comparison practices (on the top) and the differences in Medicare expenditures between non-CPC+ and non-comparison practices (on the bottom). Solid lines show the 2019 values, which we repeat twice within the same graph to contrast the 2019 values against values in both 2020 and 2021. Dashed lines show the 2020 and 2021 values. We used a concentration weight constructed at the state-HRR level such that non-CPC+ practices had the same level of representation (in terms of beneficiary months) as CPC+ practices in the same state and HRR and SSP group, and non-comparison practices had the same level of representation as comparison practices in the same state and HRR and SSP group. Non-CPC+ and non-comparison practices were not matched or weighted by beneficiary or practice characteristics.

CPC+ = Comprehensive Primary Care Plus; HRR = hospital referral region; SSP = Medicare Shared Savings Program.

Figure 5.D.viii. Growth in UCC visits was larger for non-CPC+ practices than for non-comparison practices during the pandemic period (relative to 2019)

Panel A. Track 1





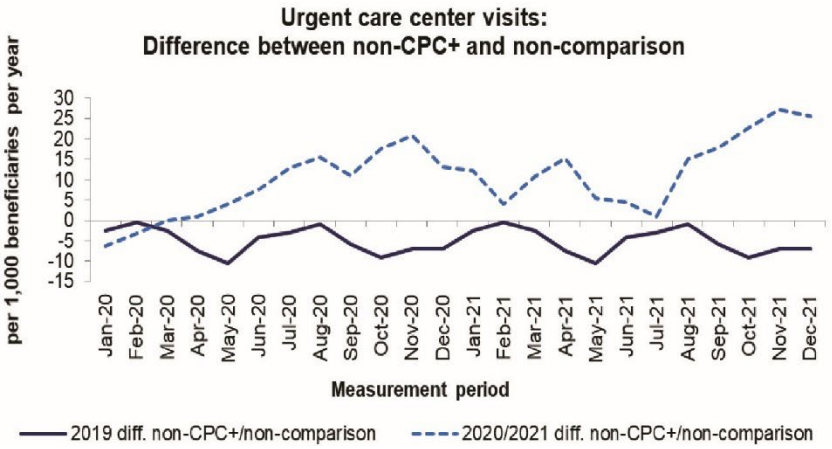
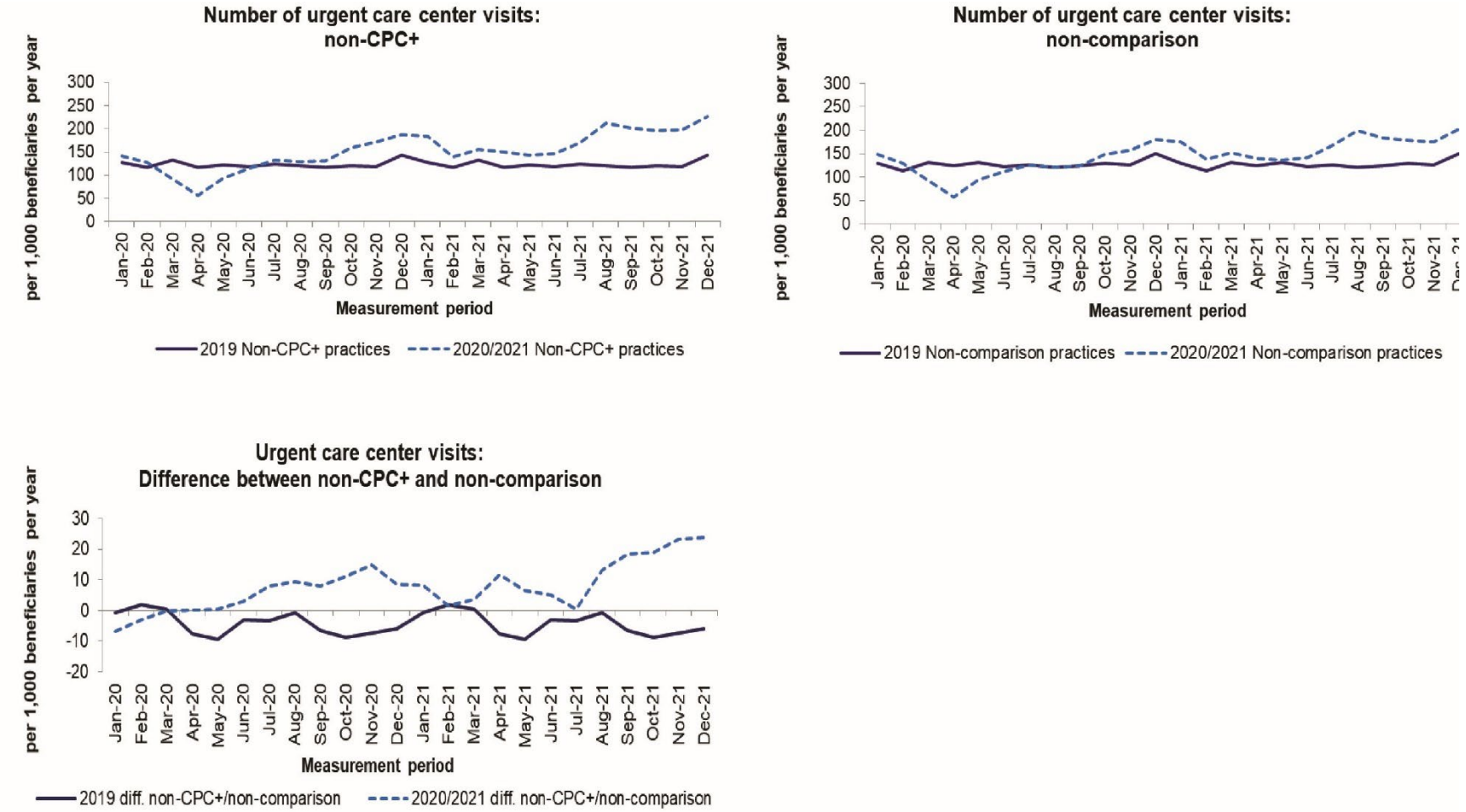


Figure 5.D.viii. ***(continued)***

Panel B. Track 2



Source: Mathematica's analysis of Medicare Part A and B claims data from January 2019 through December 2021.

Note: Lines show UCC visits for non-CPC+ and non-comparison practices (on the top) and the differences in UCC visits between non-CPC+ and non-comparison practices (on the bottom). Solid lines show the 2019 values, which we repeat twice within the same graph to contrast the 2019 values against values in both 2020 and 2021. Dashed lines show the 2020 and 2021 values. We used a concentration weight constructed at the state-HRR level such that non-CPC+ practices had the same level of representation (in terms of beneficiary months) as CPC+ practices in the same state and HRR and SSP group, and non-comparison practices had the same level of representation as comparison practices in the same state and HRR and SSP group. Non-CPC+ and non-comparison practices were not matched or weighted by beneficiary or practice characteristics.

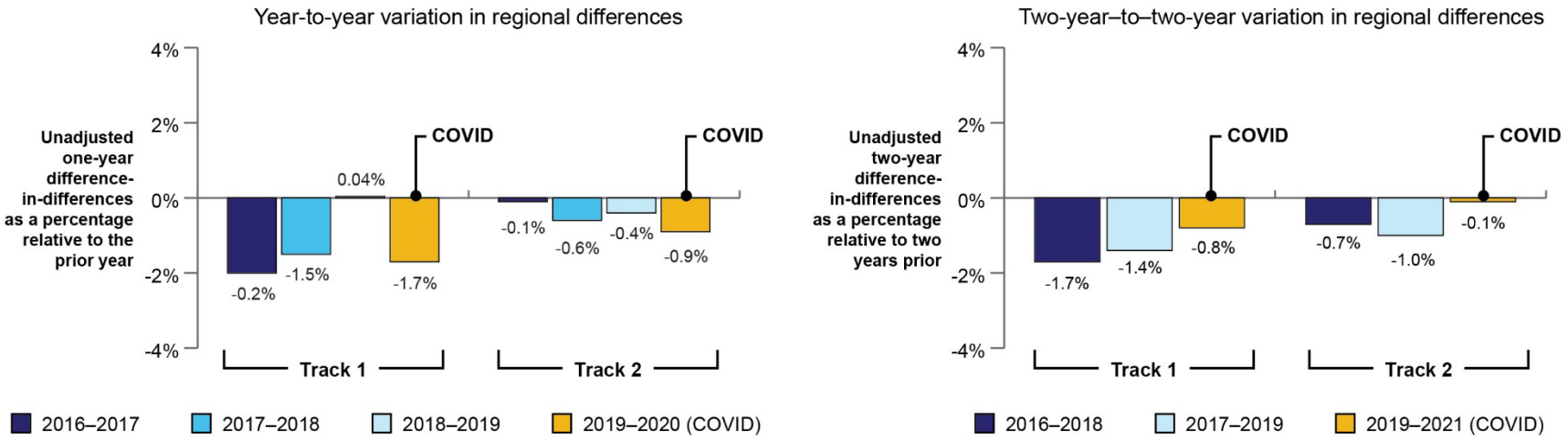
CPC+ = Comprehensive Primary Care Plus; HRR = hospital referral region; SSP = Medicare Shared Savings Program; UCC = urgent care center.

B. Year-to-year and two-year–to–two-year variation in acute care utilization and expenditures

The unadjusted difference between non-CPC+ and non-comparison practices in year-to-year and two-year–to–two-year variation (in percentage terms) in acute care utilization and expenditures for acute hospitalizations from 2019 to 2021 were not substantially greater than the historical variation that occurred before the pandemic from 2016 to 2019 (Figure 5.D.ix). For some outcomes in some years, the differences in the year-to-year and two-year–to–two-year variation were noticeable (both before and during the pandemic), but for all outcomes, the difference was within a range of –3 to 2 percent. The one exception is UCC visits, where growth during the pandemic was substantial relative to historical variation.

Figure 5.D.ix. Year-to-year and two-year–to–two-year variation among non-CPC+ practices relative to non-comparison practices, in which changes in the regional differences were not substantially greater than historical changes, except for UCC visits

Panel A. Acute care hospitalizations



Panel B. Acute inpatient expenditures

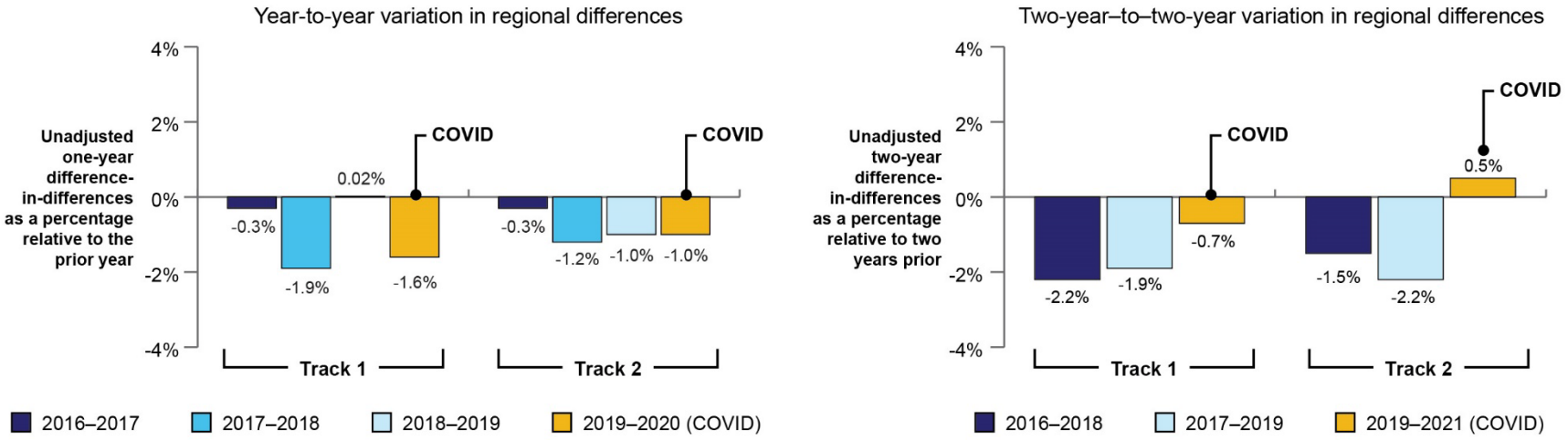
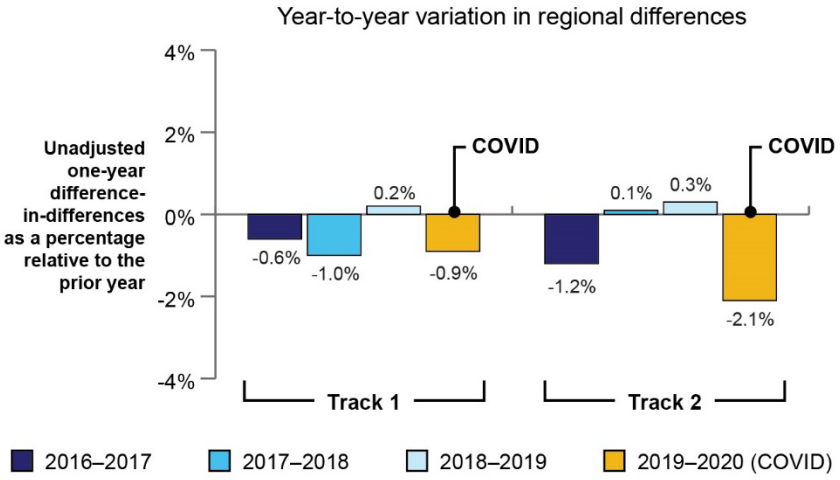
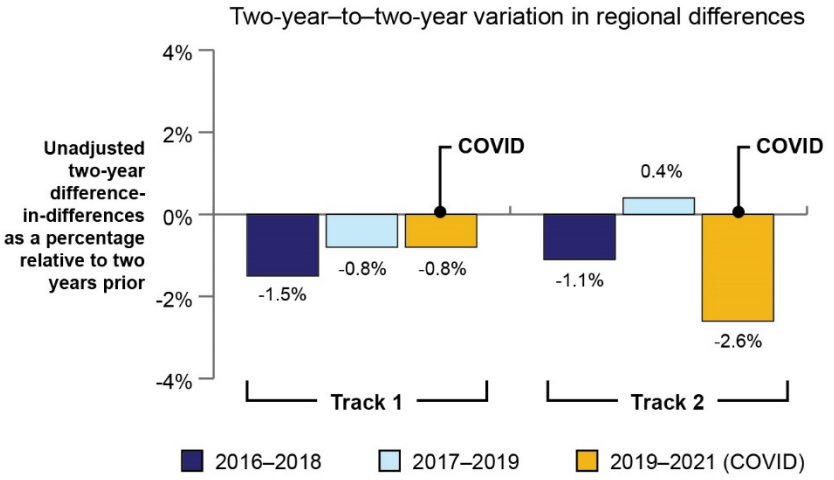


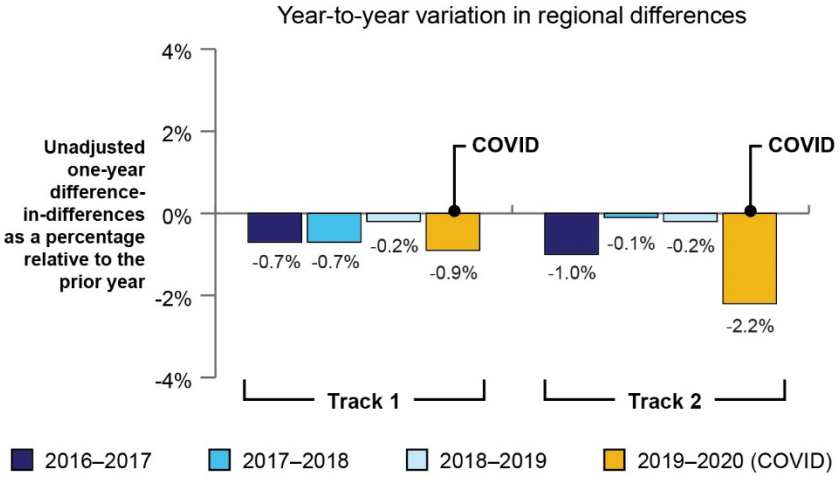
Figure 5.D.ix. ***(continued)***

Panel C. Outpatient ED visits





Panel D. Outpatient ED expenditures



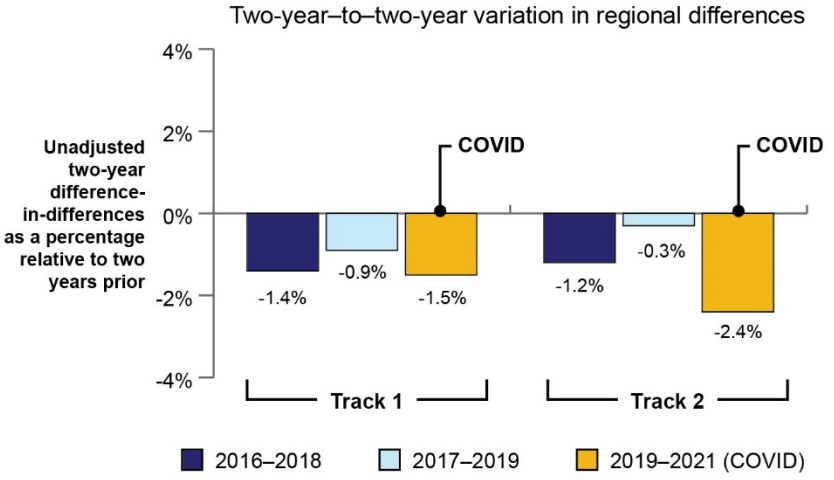
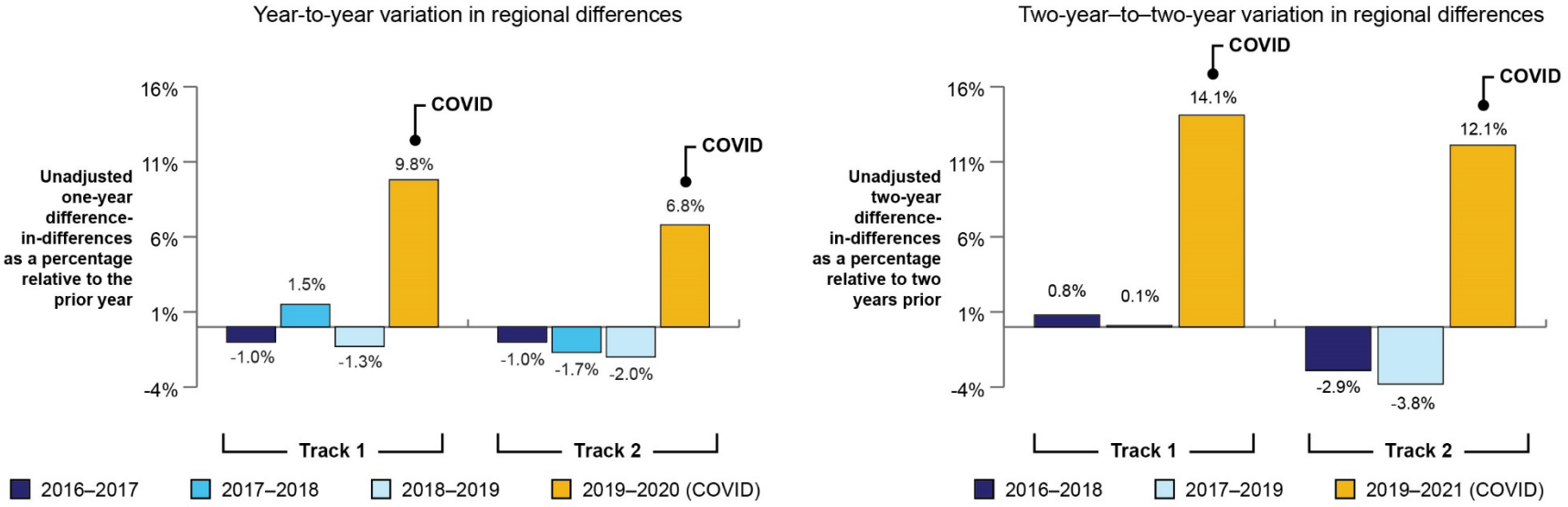


Figure 5.D.ix. ***(continued)***

Panel E. UCC visits



Source: Mathematica's analysis of Medicare claims data from January 2016 through December 2021.

Note: The figures show the unadjusted year-to-year and two-year–to–two-year variation (in percentage terms) in the differences of key impact outcomes between non-CPC+ and non-comparison practices from 2016 and 2021. In each panel, we compared the variation during the pandemic period (that is, bars with blue shading and having "COVID" in their legend labels) with pre-pandemic, historical variation (that is, bars with red shading). For year-to-year variation, we do not include the 2020–2021 variation in the post-pandemic period because none of our regression analyses is based on this year-to-year variation. For two-year–to–two-year variation, we omitted the 2018–2020 variation in the pre-pandemic period because it contains a year during the pandemic period and is, therefore, less helpful to serve as a historical benchmark for the 2019-to-2021 change. For these practices, we used a concentration weight constructed at the state-HRR level such that non-CPC+ practices had the same level of representation (in terms of beneficiary months) as CPC+ practices in the same state and HRR and SSP group, and non-comparison practices had the same level of representation as comparison practices in the same state and HRR and SSP group. We winsorized the weights at the 99th percentile.

CPC+ = Comprehensive Primary Care Plus; ED = emergency department; HRR = hospital referral region; SSP = Medicare Shared Savings Program; UCC = urgent care center.

C. Regional differences among *selected* practices in health care utilization and expenditures, when including COVID-19-related control variables

In the main text, we showed the estimated differences between non-CPC+ and non-comparison practices in health care utilization and expenditures from 2019 and 2021, before and after including the COVID-19-related controls (Tables 5.D.6 and 5.D.7). Here, in Tables 5.D.v and 5.D.vi, we show the corresponding results for CPC+ and comparison practices. Consistent with the findings for non-CPC+ and non-comparison practices, the COVID-19-related controls typically reduced the differences in 2020 but had mixed effects in 2021 for CPC+ and comparison practices.

Table 5.D.v. Using COVID-19-related controls typically reduced or had little effect on the regional differences among selected practices in 2020, but the pattern was mixed in 2021 (Track 1)

	Track 1 All, Unadjusted means		Track 1 All, Regression-adjusted CPC+ vs. comparison differences relative to 2019		Track 1 SSP, Unadjusted means		Track 1 SSP, Regression-adjusted CPC+ vs. comparison differences relative to 2019		Track 1 Non-SSP, Unadjusted means		Track 1 Non-SSP, Regression-adjusted CPC+ vs. comparison differences relative to 2019	
	CPC+	Comparison	Estimate (SE)	Percentage difference relative to 2019 comparison mean[a]	CPC+	Comparison	Estimate (SE)	Percentage difference relative to 2019 comparison mean[a]	CPC+	Comparison	Estimate (SE)	Percentage difference relative to 2019 comparison mean[a]
Medicare Part A and B expenditures without enhanced payments for CPC+ and SSP												
Using practice fixed effects and beneficiary controls												
2019 (Jan–Dec)	$990	$992	NA	NA	$1,012	$1,019	NA	NA	$966	$961	NA	NA
2020 (Jan–Dec)	$946	$959	-$8.2** ($3.4)	-0.8%	$966	$979	-$2.4 ($4.9)	-0.2%	$925	$936	-$12.8*** ($4.9)	-1.3%
2021 (Jan–Dec)	$1,037	$1,053	-$11.3*** ($3.9)	-1.1%	$1,069	$1,080	-$4.2 ($5.3)	-0.4%	$1,004	$1,022	-$16.6*** ($5.7)	-1.7%
2020–2021 (combined COVID years)	$991	$1,006	-$9.7*** ($3.3)	-1.0%	$1,017	$1,030	-$3.3 ($4.5)	-0.3%	$965	$979	-$14.7*** ($4.7)	-1.5%
Using practice fixed effects, beneficiary controls, and COVID-19-related controls												
2019 (Jan–Dec)	$990	$992	NA	NA	$1,012	$1,019	NA	NA	$966	$961	NA	NA
2020 (Jan–Dec)	$946	$959	-$4.2 ($3.6)	-0.4%	$966	$979	$2.6 ($5.2)	0.3%	$925	$936	-$9.2* ($4.9)	-1.0%
2021 (Jan–Dec)	$1,037	$1,053	-$8.7** ($4.0)	-0.9%	$1,069	$1,080	-$6.7 ($5.8)	-0.7%	$1,004	$1,022	-$11.8** ($5.6)	-1.2%
2020–2021 (combined COVID years)	$991	$1,006	-$6.5* ($3.3)	-0.7%	$1,017	$1,030	-$1.9 ($4.8)	-0.2%	$965	$979	-$10.6** ($4.7)	-1.1%
Acute hospitalizations (short-stay acute care and critical access hospitals)												
Using practice fixed effects and beneficiary controls												
2019 (Jan–Dec)	283	284	NA	NA	285	288	NA	NA	282	280	NA	NA
2020 (Jan–Dec)	243	246	-1.1 (1.3)	-0.4%	245	249	0.9 (1.7)	0.3%	241	242	-2.7 (1.9)	-1.0%
2021 (Jan–Dec)	244	246	-1.3 (1.4)	-0.5%	249	252	0.5 (1.9)	0.2%	238	240	-2.6 (2.1)	-0.9%
2020–2021 (combined COVID years)	243	246	-1.2 (1.2)	-0.4%	247	250	0.7 (1.6)	0.2%	240	241	-2.7 (1.8)	-1.0%

Table 5.D.v. ***(continued)***

	Track 1 All, Unadjusted means		Track 1 All, Regression-adjusted CPC+ vs. comparison differences relative to 2019		Track 1 SSP, Unadjusted means		Track 1 SSP, Regression-adjusted CPC+ vs. comparison differences relative to 2019		Track 1 Non-SSP, Unadjusted means		Track 1 Non-SSP, Regression-adjusted CPC+ vs. comparison differences relative to 2019	
	CPC+	Comparison	Estimate (SE)	Percentage difference relative to 2019 comparison mean[a]	CPC+	Comparison	Estimate (SE)	Percentage difference relative to 2019 comparison mean[a]	CPC+	Comparison	Estimate (SE)	Percentage difference relative to 2019 comparison mean[a]
Using practice fixed effects, beneficiary controls, and COVID-19-related controls												
2019 (Jan–Dec)	283	284	NA	NA	285	288	NA	NA	282	280	NA	NA
2020 (Jan–Dec)	243	246	-0.8 (1.4)	-0.3%	245	249	0.8 (1.9)	0.3%	241	242	-1.7 (2.0)	-0.6%
2021 (Jan–Dec)	244	246	-0.4 (1.5)	-0.2%	249	252	1.1 (2.0)	0.4%	238	240	-1.1 (2.1)	-0.4%
2020–2021 (combined COVID years)	243	246	-0.6 (1.2)	-0.2%	247	250	0.9 (1.7)	0.3%	240	241	-1.3 (1.8)	-0.5%
Expenditures on acute hospitalizations (short-stay acute care and critical access hospitals) (per beneficiary per month)												
Using practice fixed effects and beneficiary controls												
2019 (Jan–Dec)	$294	$306	NA	NA	$302	$313	NA	NA	$286	$298	NA	NA
2020 (Jan–Dec)	$279	$296	-$3.7* ($1.9)	-1.2%	$286	$300	-$0.5 ($2.7)	-0.2%	$272	$290	-$6.0** ($2.6)	-2.0%
2021 (Jan–Dec)	$293	$309	-$3.9** ($2.0)	-1.3%	$305	$317	-$0.7 ($2.8)	-0.2%	$282	$301	-$6.5** ($2.9)	-2.2%
2020–2021 (combined COVID years)	$286	$303	-$3.8** ($1.7)	-1.2%	$295	$309	-$0.6 ($2.4)	-0.2%	$277	$295	-$6.3*** ($2.4)	-2.1%
Using practice fixed effects, beneficiary controls, and COVID-19-related controls												
2019 (Jan–Dec)	$294	$306	NA	NA	$302	$313	NA	NA	$286	$298	NA	NA
2020 (Jan–Dec)	$279	$296	-$2.3 ($2.0)	-0.7%	$286	$300	$0.1 ($2.9)	0.04%	$272	$290	-$3.0 ($2.8)	-1.0%
2021 (Jan–Dec)	$293	$309	-$1.9 ($2.1)	-0.6%	$305	$317	-$0.4 ($3.0)	-0.1%	$282	$301	-$3.8 ($3.0)	-1.3%
2020–2021 (combined COVID years)	$286	$303	-$2.1 ($1.7)	-0.7%	$295	$309	-$0.1 ($2.5)	-0.03%	$277	$295	-$3.4 ($2.5)	-1.2%

Table 5.D.v. ***(continued)***

	Track 1 All, Unadjusted means		Track 1 All, Regression-adjusted CPC+ vs. comparison differences relative to 2019		Track 1 SSP, Unadjusted means		Track 1 SSP, Regression-adjusted CPC+ vs. comparison differences relative to 2019		Track 1 Non-SSP, Unadjusted means		Track 1 Non-SSP, Regression-adjusted CPC+ vs. comparison differences relative to 2019	
	CPC+	Comparison	Estimate (SE)	Percentage difference relative to 2019 comparison mean[a]	CPC+	Comparison	Estimate (SE)	Percentage difference relative to 2019 comparison mean[a]	CPC+	Comparison	Estimate (SE)	Percentage difference relative to 2019 comparison mean[a]
Outpatient ED visits, including observation stays												
Using practice fixed effects and beneficiary controls												
2019 (Jan–Dec)	485	494	NA	NA	470	475	NA	NA	500	515	NA	NA
2020 (Jan–Dec)	378	389	-0.5 (2.3)	-0.1%	362	372	-2.0 (3.0)	-0.4%	393	407	1.7 (3.4)	0.3%
2021 (Jan–Dec)	408	421	-1.9 (2.6)	-0.4%	397	406	-1.5 (3.5)	-0.3%	419	437	-1.9 (3.9)	-0.4%
2020–2021 (combined COVID years)	393	405	-1.2 (2.3)	-0.2%	380	390	-1.7 (3.0)	-0.4%	406	422	-0.1 (3.4)	-0.02%
Using practice fixed effects, beneficiary controls, and COVID-19-related controls												
2019 (Jan–Dec)	485	494	NA	NA	470	475	NA	NA	500	515	NA	NA
2020 (Jan–Dec)	378	389	-0.6 (2.3)	-0.1%	362	372	-0.4 (3.1)	-0.1%	393	407	0.0 (3.6)	0.00%
2021 (Jan–Dec)	408	421	-4.4* (2.7)	-0.9%	397	406	-4.1 (3.6)	-0.9%	419	437	-4.8 (3.8)	-0.9%
2020–2021 (combined COVID years)	393	405	-2.6 (2.3)	-0.5%	380	390	-2.2 (3.0)	-0.5%	406	422	-2.5 (3.4)	-0.5%
Expenditures on outpatient ED visits, including observation stays (per beneficiary per month)												
Using practice fixed effects and beneficiary controls												
2019 (Jan–Dec)	$29	$30	NA	NA	$29	$29	NA	NA	$30	$31	NA	NA
2020 (Jan–Dec)	$24	$26	-$0.1 ($0.2)	-0.3%	$23	$25	-$0.1 ($0.3)	-0.2%	$25	$27	-$0.2 ($0.3)	-0.5%
2021 (Jan–Dec)	$28	$30	-$0.2 ($0.3)	-0.7%	$27	$29	-$0.2 ($0.3)	-0.7%	$29	$31	-$0.2 ($0.4)	-0.7%
2020–2021 (combined COVID years)	$26	$28	-$0.2 ($0.2)	-0.5%	$25	$27	-$0.1 ($0.3)	-0.5%	$27	$29	-$0.2 ($0.3)	-0.6%

Table 5.D.v. ***(continued)***

	Track 1 All, Unadjusted means		Track 1 All, Regression-adjusted CPC+ vs. comparison differences relative to 2019		Track 1 SSP, Unadjusted means		Track 1 SSP, Regression-adjusted CPC+ vs. comparison differences relative to 2019		Track 1 Non-SSP, Unadjusted means		Track 1 Non-SSP, Regression-adjusted CPC+ vs. comparison differences relative to 2019	
	CPC+	Comparison	Estimate (SE)	Percentage difference relative to 2019 comparison mean[a]	CPC+	Comparison	Estimate (SE)	Percentage difference relative to 2019 comparison mean[a]	CPC+	Comparison	Estimate (SE)	Percentage difference relative to 2019 comparison mean[a]
Using practice fixed effects, beneficiary controls, and COVID-19-related controls												
2019 (Jan–Dec)	$29	$30	NA	NA	$29	$29	NA	NA	$30	$31	NA	NA
2020 (Jan–Dec)	$24	$26	$0.1 ($0.2)	0.3%	$23	$25	$0.3 ($0.3)	1.1%	$25	$27	-$0.1 ($0.3)	-0.4%
2021 (Jan–Dec)	$28	$30	-$0.5* ($0.3)	-1.6%	$27	$29	-$0.4 ($0.3)	-1.5%	$29	$31	-$0.6 ($0.4)	-2.0%
2020–2021 (combined COVID years)	$26	$28	-$0.2 ($0.2)	-0.7%	$25	$27	$0.0 ($0.3)	-0.1%	$27	$29	-$0.4 ($0.3)	-1.2%
Urgent care center visits (per 1,000 beneficiaries per year)												
Using practice fixed effects and beneficiary controls												
2019 (Jan–Dec)	149	152	NA	NA	167	159	NA	NA	131	145	NA	NA
2020 (Jan–Dec)	156	144	14.5*** (3.0)	9.5%	179	152	18.0*** (3.9)	11.3%	133	135	10.9** (4.6)	7.6%
2021 (Jan–Dec)	212	201	15.4*** (4.4)	10.1%	250	219	23.5*** (6.3)	14.8%	173	181	8.1 (5.9)	5.6%
2020–2021 (combined COVID years)	184	173	14.9*** (3.4)	9.8%	214	186	20.7*** (4.7)	13.1%	153	158	9.5* (5.0)	6.6%
Using practice fixed effects, beneficiary controls, and COVID-19-related controls												
2019 (Jan–Dec)	149	152	NA	NA	167	159	NA	NA	131	145	NA	NA
2020 (Jan–Dec)	156	144	14.0*** (2.9)	9.2%	179	152	15.5*** (3.9)	9.7%	133	135	12.8*** (4.5)	8.8%
2021 (Jan–Dec)	212	201	3.8 (4.3)	2.5%	250	219	-2.1 (6.6)	-1.3%	173	181	7.6 (5.6)	5.3%
2020–2021 (combined COVID years)	184	173	8.8*** (3.2)	5.8%	214	186	7.1 (4.3)	4.5%	153	158	10.0** (4.6)	6.9%

Table 5.D.v. ***(continued)***

	Track 1 All, Unadjusted means		Track 1 All, Regression-adjusted CPC+ vs. comparison differences relative to 2019		Track 1 SSP, Unadjusted means		Track 1 SSP, Regression-adjusted CPC+ vs. comparison differences relative to 2019		Track 1 Non-SSP, Unadjusted means		Track 1 Non-SSP, Regression-adjusted CPC+ vs. comparison differences relative to 2019	
	CPC+	Comparison	Estimate (SE)	Percentage difference relative to 2019 comparison mean[a]	CPC+	Comparison	Estimate (SE)	Percentage difference relative to 2019 comparison mean[a]	CPC+	Comparison	Estimate (SE)	Percentage difference relative to 2019 comparison mean[a]
Unweighted sample sizes												
Number of practices	1,373	5,242			738	2,979			635	2,263		
Average number of beneficiaries per month	1,020,196	3,529,430			516,727	2,063,138			503,469	1,466,292		

Source: Mathematica's analysis of Medicare claims data from January 2019 through December 2021.

Note: Estimates in the table are derived from separate models run at the practice-month-year level that are regression-adjusted for (1) baseline beneficiary characteristics and practice fixed effects and (2) baseline beneficiary characteristics, practice fixed effects, and COVID-19-related controls. For CPC+ practices, observations were weighted by the number of Medicare FFS beneficiaries assigned to the practice during the month and year. For comparison practices, the weight is a product of the number of assigned beneficiaries and the matching weight. Standard errors are clustered at the practice level. For a detailed description of methods, see Supplement 5.

*/**/*** Significantly different from zero at the 0.10/0.05/0.01 levels, two-tailed test.

[a] To calculate these percentages, we divided the estimated CPC+ versus comparison differences by the unadjusted 2019 comparison mean for the outcome.

CPC+ = Comprehensive Primary Care Plus; ED = emergency department; FFS = fee-for-service; HCC = hierarchical condition category; HRR = hospital referral region; NA = not applicable; SE = standard error; SSP = Medicare Shared Savings Program.

Table 5.D.vi. Using COVID-19-related controls typically reduced or had little effect on the regional differences among selected practices in 2020, but the pattern was mixed in 2021 (Track 2)

	Track 2 All, Unadjusted means		Track 2 All, Regression-adjusted CPC+ vs. comparison differences relative to 2019		Track 2 SSP, Unadjusted means		Track 2 SSP, Regression-adjusted CPC+ vs. comparison differences relative to 2019		Track 2 Non-SSP, Unadjusted means		Track 2 Non-SSP, Regression-adjusted CPC+ vs. comparison differences relative to 2019	
	CPC+	Comparison	Estimate (SE)	Percentage difference relative to 2019 comparison mean[a]	CPC+	Comparison	CPC+	Comparison	Estimate (SE)	Percentage difference relative to 2019 comparison mean[a]	CPC+	Comparison
Medicare Part A and B expenditures without enhanced payments for CPC+ and SSP												
Using practice fixed effects and beneficiary controls												
2019 (Jan–Dec)	$986	$992	NA	NA	$1,005	$1,016	NA	NA	$971	$973	NA	NA
2020 (Jan–Dec)	$944	$958	-$7.1** ($3.4)	-0.7%	$953	$980	-$14.2*** ($4.9)	-1.4%	$936	$940	-$3.3 ($4.6)	-0.3%
2021 (Jan–Dec)	$1,031	$1,043	-$5.9 ($3.9)	-0.6%	$1,045	$1,069	-$13.9** ($5.7)	-1.4%	$1,019	$1,022	-$2.4 ($5.5)	-0.3%
2020–2021 (combined COVID years)	$987	$1,000	-$6.5** ($3.2)	-0.7%	$999	$1,025	-$14.1*** ($4.7)	-1.4%	$978	$981	-$2.9 ($4.5)	-0.3%
Using practice fixed effects, beneficiary controls, and COVID-19-related controls												
2019 (Jan–Dec)	$986	$992	NA	NA	$1,005	$1,016	NA	NA	$971	$973	NA	NA
2020 (Jan–Dec)	$944	$958	-$3.3 ($3.5)	-0.3%	$953	$980	-$11.1** ($5.3)	-1.1%	$936	$940	$1.6 ($4.6)	0.2%
2021 (Jan–Dec)	$1,031	$1,043	-$2.9 ($4.0)	-0.3%	$1,045	$1,069	-$15.2*** ($5.8)	-1.5%	$1,019	$1,022	$4.0 ($5.5)	0.4%
2020–2021 (combined COVID years)	$987	$1,000	-$3.1 ($3.3)	-0.3%	$999	$1,025	-$13.2*** ($4.8)	-1.3%	$978	$981	$2.8 ($4.5)	0.3%
Acute hospitalizations (short-stay acute care and critical access hospitals)												
Using practice fixed effects and beneficiary controls												
2019 (Jan–Dec)	286	286	NA	NA	296	290	NA	NA	278	284	NA	NA
2020 (Jan–Dec)	245	247	-1.1 (1.3)	-0.4%	253	251	-4.4** (2.0)	-1.5%	239	243	0.7 (1.7)	0.2%
2021 (Jan–Dec)	246	245	0.6 (1.4)	0.2%	256	251	-1.7 (2.1)	-0.6%	237	241	1.4 (1.9)	0.5%
2020–2021 (combined COVID years)	245	246	-0.3 (1.2)	-0.1%	255	251	-3.1* (1.9)	-1.1%	238	242	1.0 (1.6)	0.4%

Table 5.D.vi. ***(continued)***

	Track 2 All, Unadjusted means		Track 2 All, Regression-adjusted CPC+ vs. comparison differences relative to 2019		Track 2 SSP, Unadjusted means		Track 2 SSP, Regression-adjusted CPC+ vs. comparison differences relative to 2019		Track 2 Non-SSP, Unadjusted means		Track 2 Non-SSP, Regression-adjusted CPC+ vs. comparison differences relative to 2019	
	CPC+	Comparison	Estimate (SE)	Percentage difference relative to 2019 comparison mean[a]	CPC+	Comparison	CPC+	Comparison	Estimate (SE)	Percentage difference relative to 2019 comparison mean[a]	CPC+	Comparison
Using practice fixed effects, beneficiary controls, and COVID-19-related controls												
2019 (Jan–Dec)	286	286	NA	NA	296	290	NA	NA	278	284	NA	NA
2020 (Jan–Dec)	245	247	-0.7 (1.4)	-0.2%	253	251	-4.0* (2.1)	-1.4%	239	243	0.9 (1.8)	0.3%
2021 (Jan–Dec)	246	245	0.5 (1.4)	0.2%	256	251	-2.6 (2.1)	-0.9%	237	241	2.2 (1.9)	0.8%
2020–2021 (combined COVID years)	245	246	-0.1 (1.2)	-0.03%	255	251	-3.3* (1.9)	-1.1%	238	242	1.5 (1.6)	0.6%
Expenditures on acute hospitalizations (short-stay acute care and critical access hospitals) (per beneficiary per month)												
Using practice fixed effects and beneficiary controls												
2019 (Jan–Dec)	$298	$309	NA	NA	$306	$316	NA	NA	$292	$303	NA	NA
2020 (Jan–Dec)	$284	$297	-$2.1 ($2.0)	-0.7%	$290	$305	-$3.6 ($3.0)	-1.1%	$279	$291	-$1.5 ($2.7)	-0.5%
2021 (Jan–Dec)	$297	$308	$0.2 ($2.1)	0.1%	$305	$316	-$0.8 ($3.1)	-0.2%	$291	$301	-$0.4 ($3.0)	-0.1%
2020–2021 (combined COVID years)	$290	$302	-$0.9 ($1.8)	-0.3%	$298	$310	-$2.2 ($2.6)	-0.7%	$285	$296	-$0.9 ($2.5)	-0.3%
Using practice fixed effects, beneficiary controls, and COVID-19-related controls												
2019 (Jan–Dec)	$298	$309	NA	NA	$306	$316	NA	NA	$292	$303	NA	NA
2020 (Jan–Dec)	$284	$297	$0.3 ($2.1)	0.1%	$290	$305	-$1.7 ($3.1)	-0.5%	$279	$291	$2.1 ($2.7)	0.7%
2021 (Jan–Dec)	$297	$308	$1.5 ($2.2)	0.5%	$305	$316	-$2.0 ($3.2)	-0.6%	$291	$301	$3.3 ($3.0)	1.1%
2020–2021 (combined COVID years)	$290	$302	$0.9 ($1.8)	0.3%	$298	$310	-$1.9 ($2.7)	-0.6%	$285	$296	$2.7 ($2.5)	0.9%

Table 5.D.vi. ***(continued)***

	Track 2 All, Unadjusted means		Track 2 All, Regression-adjusted CPC+ vs. comparison differences relative to 2019		Track 2 SSP, Unadjusted means		Track 2 SSP, Regression-adjusted CPC+ vs. comparison differences relative to 2019		Track 2 Non-SSP, Unadjusted means		Track 2 Non-SSP, Regression-adjusted CPC+ vs. comparison differences relative to 2019	
	CPC+	Comparison	Estimate (SE)	Percentage difference relative to 2019 comparison mean[a]	CPC+	Comparison	CPC+	Comparison	Estimate (SE)	Percentage difference relative to 2019 comparison mean[a]	CPC+	Comparison
Outpatient ED visits, including observation stays												
Using practice fixed effects and beneficiary controls												
2019 (Jan–Dec)	486	488	NA	NA	471	472	NA	NA	497	501	NA	NA
2020 (Jan–Dec)	380	385	-1.6 (2.1)	-0.3%	364	376	-10.2*** (3.0)	-2.2%	393	392	6.2** (2.8)	1.2%
2021 (Jan–Dec)	408	415	-3.4 (2.4)	-0.7%	393	408	-14.1*** (3.4)	-3.0%	420	421	6.4* (3.4)	1.3%
2020–2021 (combined COVID years)	394	400	-2.5 (2.1)	-0.5%	378	392	-12.2*** (3.0)	-2.6%	406	407	6.3** (2.8)	1.3%
Using practice fixed effects, beneficiary controls, and COVID-19-related controls												
2019 (Jan–Dec)	486	488	NA	NA	471	472	NA	NA	497	501	NA	NA
2020 (Jan–Dec)	380	385	-1.3 (2.2)	-0.3%	364	376	-10.5*** (3.1)	-2.2%	393	392	8.3*** (2.9)	1.7%
2021 (Jan–Dec)	408	415	-3.6 (2.5)	-0.7%	393	408	-14.9*** (3.5)	-3.2%	420	421	7.9** (3.3)	1.6%
2020–2021 (combined COVID years)	394	400	-2.5 (2.2)	-0.5%	378	392	-12.8*** (3.0)	-2.7%	406	407	8.1*** (2.8)	1.6%
Expenditures on outpatient ED visits, including observation stays (per beneficiary per month)												
Using practice fixed effects and beneficiary controls												
2019 (Jan–Dec)	$29	$30	NA	NA	$29	$31	NA	NA	$30	$30	NA	NA
2020 (Jan–Dec)	$24	$26	-$0.2 ($0.2)	-0.8%	$23	$26	-$0.3 ($0.3)	-1.1%	$25	$25	-$0.2 ($0.3)	-0.5%
2021 (Jan–Dec)	$28	$30	-$0.5** ($0.3)	-1.7%	$27	$30	-$1.2*** ($0.4)	-3.9%	$29	$29	$0.1 ($0.3)	0.4%
2020–2021 (combined COVID years)	$26	$28	-$0.4* ($0.2)	-1.2%	$25	$28	-$0.8** ($0.3)	-2.5%	$27	$27	$0.0 ($0.3)	-0.03%

Table 5.D.vi. ***(continued)***

	Track 2 All, Unadjusted means		Track 2 All, Regression-adjusted CPC+ vs. comparison differences relative to 2019		Track 2 SSP, Unadjusted means		Track 2 SSP, Regression-adjusted CPC+ vs. comparison differences relative to 2019		Track 2 Non-SSP, Unadjusted means		Track 2 Non-SSP, Regression-adjusted CPC+ vs. comparison differences relative to 2019	
	CPC+	Comparison	Estimate (SE)	Percentage difference relative to 2019 comparison mean[a]	CPC+	Comparison	CPC+	Comparison	Estimate (SE)	Percentage difference relative to 2019 comparison mean[a]	CPC+	Comparison
Using practice fixed effects, beneficiary controls, and COVID-19-related controls												
2019 (Jan–Dec)	$29	$30	NA	NA	$29	$31	NA	NA	$30	$30	NA	NA
2020 (Jan–Dec)	$24	$26	-$0.2 ($0.2)	-0.7%	$23	$26	-$0.3 ($0.3)	-1.0%	$25	$25	-$0.1 ($0.3)	-0.5%
2021 (Jan–Dec)	$28	$30	-$0.7*** ($0.3)	-2.2%	$27	$30	-$1.2*** ($0.4)	-3.8%	$29	$29	-$0.1 ($0.3)	-0.4%
2020–2021 (combined COVID years)	$26	$28	-$0.4** ($0.2)	-1.5%	$25	$28	-$0.7** ($0.3)	-2.4%	$27	$27	-$0.1 ($0.3)	-0.4%
Urgent care center visits (per 1,000 beneficiaries per year)												
Using practice fixed effects and beneficiary controls												
2019 (Jan–Dec)	134	145	NA	NA	138	146	NA	NA	131	145	NA	NA
2020 (Jan–Dec)	135	140	7.9*** (2.8)	5.5%	141	142	7.1* (4.1)	4.9%	131	138	8.3** (3.8)	5.7%
2021 (Jan–Dec)	186	188	11.4** (4.7)	7.9%	195	197	5.6 (7.5)	3.8%	179	181	15.9*** (6.0)	11.0%
2020–2021 (combined COVID years)	161	164	9.7*** (3.5)	6.7%	168	170	6.3 (5.4)	4.4%	155	159	12.1*** (4.6)	8.3%
Using practice fixed effects, beneficiary controls, and COVID-19-related controls												
2019 (Jan–Dec)	134	145	NA	NA	138	146	NA	NA	131	145	NA	NA
2020 (Jan–Dec)	135	140	8.1*** (2.8)	5.6%	141	142	4.3 (4.5)	2.9%	131	138	10.5*** (3.7)	7.3%
2021 (Jan–Dec)	186	188	7.0 (4.5)	4.8%	195	197	-6.2 (7.7)	-4.3%	179	181	14.3** (5.7)	9.8%
2020–2021 (combined COVID years)	161	164	7.5** (3.4)	5.2%	168	170	-1.1 (5.5)	-0.7%	155	159	12.4*** (4.4)	8.6%

Table 5.D.vi. ***(continued)***

	Track 2 All, Unadjusted means		Track 2 All, Regression-adjusted CPC+ vs. comparison differences relative to 2019		Track 2 SSP, Unadjusted means		Track 2 SSP, Regression-adjusted CPC+ vs. comparison differences relative to 2019		Track 2 Non-SSP, Unadjusted means		Track 2 Non-SSP, Regression-adjusted CPC+ vs. comparison differences relative to 2019	
	CPC+	Comparison	Estimate (SE)	Percentage difference relative to 2019 comparison mean[a]	CPC+	Comparison	CPC+	Comparison	Estimate (SE)	Percentage difference relative to 2019 comparison mean[a]	CPC+	Comparison
Unweighted sample sizes												
Number of practices	1,515	3,783			636	1,817			879	1,966		
Average number of beneficiaries per month	1,249,560	2,984,268			554,430	1,491,846			695,130	1,492,422		

Source: Mathematica's analysis of Medicare claims data from January 2019 through December 2021.

Note: Estimates in the table are derived from separate models run at the practice-month-year level that are regression-adjusted for (1) baseline beneficiary characteristics and practice fixed effects and (2) baseline beneficiary characteristics, practice fixed effects, and COVID-19-related controls. For CPC+ practices, observations were weighted by the number of Medicare FFS beneficiaries assigned to the practice during the month and year. For comparison practices, the weight is a product of the number of assigned beneficiaries and the matching weight. Standard errors are clustered at the practice level. For a detailed description of methods, see Supplement 5.

*/**/*** Significantly different from zero at the 0.10/0.05/0.01 levels, two-tailed test.

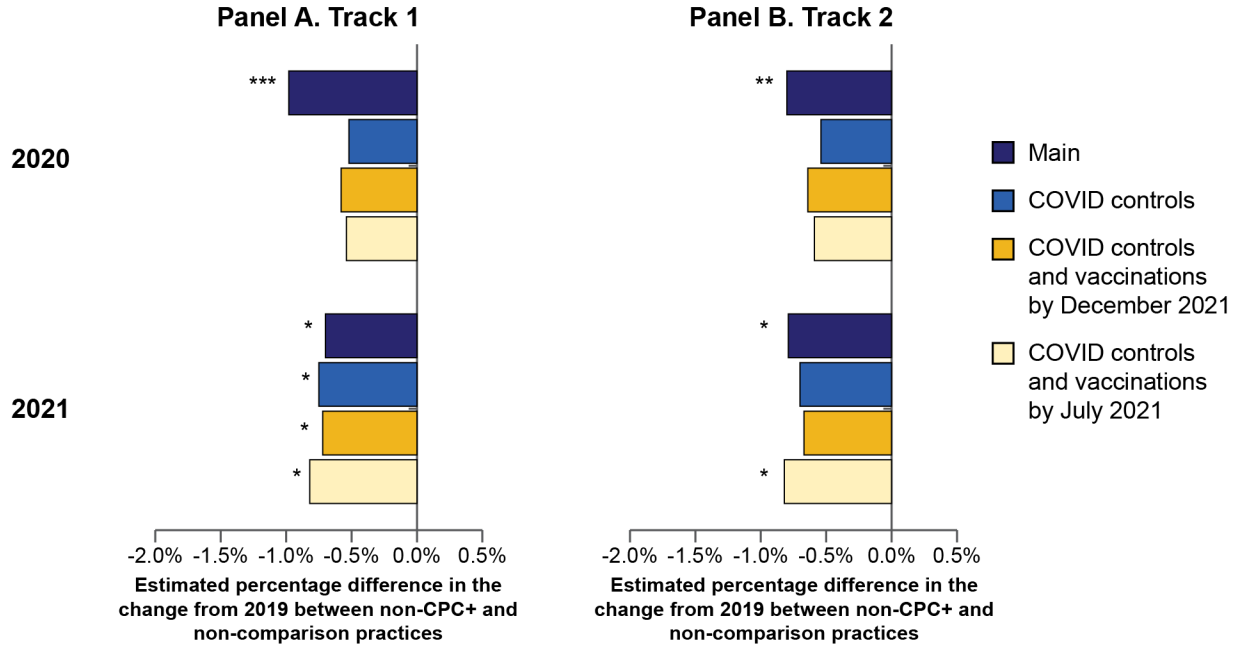
[a] To calculate these percentages, we divided the estimated CPC+ versus comparison differences by the unadjusted 2019 comparison mean for the outcome.

CPC+ = Comprehensive Primary Care Plus; ED = emergency department; FFS = fee-for-service; HCC = hierarchical condition category; HRR = hospital referral region; NA = not applicable; SE = standard error; SSP = Medicare Shared Savings Program.

D. Sensitivity tests controlling for COVID-19 vaccination rates

As described in Section 5.D.4 of Appendix 5.D, differences in COVID-19 vaccination take-up among the 65+ population could lead to differential changes in health utilization and expenditures from 2019 to 2021 between CPC+ and comparison regions. To shed light on whether our results on the total effects of COVID-19 were substantially influenced by COVID-19 vaccination take-up, we ran sensitivity checks that included county-level COVID-19 vaccination rates as additional regional COVID-19-related controls in the model. Specifically, we measured the COVID-19 vaccination rate using CDC data on the percentage of a county's 65+ population being fully vaccinated against COVID-19 (that is, per CDC, having received one dose of a single-dose vaccine such as Johnson & Johnson, or two doses on different days, regardless of time interval, of either an mRNA or protein-based series such as Moderna and Pfizer). We ran separate analyses using two versions of COVID-19 vaccination rates: (1) vaccination rates as of July 2021 and (2) vaccination rates as of December 2021. Compared to our main set of COVID-controls, additionally controlling for regional COVID-19 vaccination rates had little effect on the estimated differences in Medicare expenditures (Figure 5.D.x) and the other health utilization and expenditure outcomes we studied (data not shown).

Figure 5.D.x. Controlling for COVID-19 vaccination rates had little effect on the regional differences in Medicare expenditures



Source: Mathematica's analysis of CDC county-level COVID-19 vaccination rates (by July 2021 and December 2021) and Medicare claims data (January 2019 through December 2021).

Note: Estimates are derived from separate regression models run at the practice-month-year level. "Main" are models that control for baseline beneficiary characteristics and practice fixed effects. "COVID controls" are models that control for baseline beneficiary characteristics, practice fixed effects, and COVID-19-related controls (as described in Table 5.D.5 in Section 5.D.4 of the main text). "COVID controls & vaccinations by December 2021" are "COVID controls" models that additionally account for the county-level COVID-19 vaccination rates (percentage of a county's 65+ population fully vaccinated against COVID-19, that is, if they receive one dose of a single-dose vaccine such as Johnson & Johnson, or two doses of either an mRNA or protein-based series such as Moderna and Pfizer) as of December 2021. "COVID controls & vaccinations by July 2021" are the same as "COVID controls & vaccinations by December 2021" models except that they control for the vaccination rates as of July 2021. For these practices, we used a

Figure 5.D.x. ***(continued)***

concentration weight constructed at the state-HRR level such that non-CPC+ practices had the same level of representation (in terms of beneficiary months) as CPC+ practices in the same state and HRR and SSP group, and non-comparison practices had the same level of representation as comparison practices in the same state and HRR and SSP group. We winsorized the weights at the 99th percentile. To calculate the percentage estimates, we divided the estimated non-CPC+ versus non-comparison differences by the unadjusted 2019 non-comparison mean for the outcome.

*/**/*** Significantly different from zero at the 0.10/0.05/0.01 levels, two-tailed test.

CDC = Centers for Disease Control and Prevention; CPC+ = Comprehensive Primary Care Plus; HCC = hierarchical condition category; HRR = hospital referral region; SSP = Medicare Shared Savings Program.

E. Results from a triple-differences model

An alternative option to account for regional differences due to COVID-19 is a triple-differences model, accounting for trends among unselected practices in the same regions as CPC+ and comparison practices. We explored this option in detail in the AR4 COVID analysis for changes from 2019 to 2020 (Laird et al. 2022), and we present updated information with data through 2021 here.[61]

In the triple-differences model, we account for differential regional trends by netting out the difference in changes in outcomes between non-CPC+ practices and non-comparison practices (as opposed to including COVID-19-related controls in the regression models). Introducing these additional reference groups, which likely experienced similar effects of COVID-19 as other practices within their region, the triple-differences model can remove differential changes in outcomes due to COVID-19. The underlying assumption of the triple-differences model is that, in the absence of the CPC+ model, the difference in trends between CPC+ and non-CPC+ practices would be similar to the difference in trends between comparison and non-comparison practices.

We studied how the triple-differences model accounted for the regional differences by comparing the estimates with the difference-in-differences model, both adjusted for beneficiary controls and practice fixed effects but without COVID-19-related controls. We found that:

- For Medicare expenditures, expenditures on acute hospitalizations and outpatient ED visits, and UCC visits, the triple-differences estimates were generally closer to zero than the difference-in-differences estimates for selected practices in both 2020 and 2021 for both tracks (Table 5.D.vii). For rates of acute hospitalizations and outpatient ED visits, there was not a systematic pattern between the magnitude of the triple-differences estimates and that for the difference-in-differences estimates. For example, the triple-differences estimates for rates of acute hospitalizations were closer to zero than the corresponding difference-in-differences estimates for selected practices in Track 1 in 2020, while the triple-differences estimates were larger in magnitude than the difference-in-differences estimates in 2021 and for Track 2 both years.
- Compared to the overall sample, results from the triple-differences model by SSP status were more mixed and varied by track, outcome, and year (Tables 5.D.viii and 5.D.ix).
  - For Track 1, the triple-differences estimates increased the magnitude of the regional differences estimated from the difference-in-differences model without COVID-19-related controls for selected SSP practices for all outcomes, except for UCC visits. By contrast, the triple-differences estimates reduced the magnitude of the regional differences among non-SSP practices for Medicare expenditures, rates of acute hospitalizations and expenditures for acute hospitalization, and UCC visits.

[61] See Section 5.D.4.B, Appendix 5.D, of the appendices to the Fourth Annual Report for these analyses.

  - For Track 2, we observed generally the opposite pattern from Track 1—that is, the triple-differences estimates reduced the magnitude of regional differences among SSP practices in most cases, while it generally increased the magnitude of the regional differences among non-SSP practices.

- In many of these cases, the triple-differences models changed the direction of the regional differences estimated from the difference-in-differences model without COVID-19-related controls for selected practices.

Although the triple-differences model potentially accounts for regional differences due to COVID-19, this approach has several technical limitations, as we had noted in AR4:

- **Imbalance.** First, unlike CPC+ and comparison practices, unselected practices were not matched and, therefore, were not well balanced on baseline characteristics. This lack of baseline balance requires us to make the strong assumption that any difference between unselected practices in CPC+ and comparison regions changed linearly after COVID-19. If, on average, COVID-19 affects practices with different characteristics differently, or has non-linear effects on outcomes, then the necessary parallel-trends assumption for generating unbiased impact estimates would not hold.
- **Potential spillover effects.** The triple-differences analysis assumes there are no spillovers of CPC+ on nonparticipating practices. If there are spillovers (for example, favorable impacts of CPC+ that spill over to nonparticipating practices that are owned by the same parent entity), the triple-differences model would net out part of the effect of CPC+ and dilute the estimated effects of the CPC+ model relative to estimates derived from difference-in-differences models for CPC+ and comparison practices.
- **Less power.** Finally, we likely have less power to detect effects with the triple-differences model compared to the difference-in-differences model because of the added uncertainty from estimating an additional layer of difference (Laird et al. 2022).[62]

In addition to the technical limitations described above, the triple-differences models are much more resource intensive to implement than the difference-in-differences models, meaning that using them as our "main" impact analysis approach, for multiple outcomes, would create delays for the final report work.

[62] The relevant information is in footnote 74 (page 458, Appendix 5.D of the appendices to the Fourth Annual Report [Laird et al. 2022]), which reads "The standard errors in a test run of the fourth annual report triple-differences model on Medicare expenditures were approximately 30 percent larger than the size of the standard errors in test runs of the fourth annual report difference-in-differences models on Medicare expenditures for both Track 1 and 2."

Table 5.D.vii. Triple-differences models reduced the magnitude of the differences between CPC+ and comparison practices for most outcomes.

	Track 1, CPC+ vs. comparison differences relative to 2019		Track 1, Non-CPC+ vs. non-comparison differences relative to 2019		Track 1, Triple-differences model		Track 2, CPC+ vs. comparison differences relative to 2019		Track 2, Non-CPC+ vs. non-comparison differences relative to 2019		Track 2, Triple-differences model	
	Estimate (SE)	Percentage difference relative to 2019 comparison mean[a]	Estimate (SE)	Percentage difference relative to 2019 non-comparison mean[b]	Estimate (SE)	Percentage difference relative to 2019 comparison mean[a]	Estimate (SE)	Percentage difference relative to 2019 non-comparison mean[b]	Estimate (SE)	Percentage difference relative to 2019 comparison mean[a]	Estimate (SE)	Percentage difference relative to 2019 non-comparison mean[b]
Medicare Part A and B expenditures without enhanced payments for CPC+ and SSP												
2019 (Jan–Dec)	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA
2020 (Jan–Dec)	-$8.2** ($3.4)	-0.8%	-$10.4*** ($3.7)	-1.0%	$2.2 ($5.1)	0.2%	-$7.1** ($3.4)	-0.7%	-$8.4** ($3.7)	-0.8%	$1.4 ($5.0)	0.1%
2021 (Jan–Dec)	-$11.3*** ($3.9)	-1.1%	-$7.4* ($4.3)	-0.7%	-$3.9 ($5.8)	-0.4%	-$5.9 ($3.9)	-0.6%	-$8.4* ($4.3)	-0.8%	$2.5 ($5.8)	0.3%
2020–2021 (combined COVID years)	-$9.7*** ($3.3)	-1.0%	-$8.9** ($3.5)	-0.8%	-$0.8 ($4.8)	-0.1%	-$6.5** ($3.2)	-0.7%	-$8.4** ($3.5)	-0.8%	$1.9 ($4.8)	0.2%
Acute hospitalizations (short-stay acute care and critical access hospitals)												
2019 (Jan–Dec)	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA
2020 (Jan–Dec)	-1.1 (1.3)	-0.4%	-2.9** (1.4)	-1.0%	1.8 (1.9)	0.7%	-1.1 (1.3)	-0.4%	-1.3 (1.4)	-0.4%	0.2 (1.9)	0.1%
2021 (Jan–Dec)	-1.3 (1.4)	-0.5%	1.0 (1.6)	0.3%	-2.3 (2.2)	-0.8%	0.6 (1.4)	0.2%	2.5 (1.6)	0.8%	-1.9 (2.1)	-0.7%
2020–2021 (combined COVID years)	-1.2 (1.2)	-0.4%	-1.0 (1.4)	-0.3%	-0.2 (1.8)	-0.1%	-0.3 (1.2)	-0.1%	0.5 (1.4)	0.2%	-0.8 (1.8)	-0.3%
Expenditures on acute hospitalizations (short-stay acute care and critical access hospitals) (per beneficiary per month)												
2019 (Jan–Dec)	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA
2020 (Jan–Dec)	-$3.7* ($1.9)	-1.2%	-$4.6** ($2.2)	-1.4%	$1.0 ($2.9)	0.3%	-$2.1 ($2.0)	-0.7%	-$3.0 ($2.2)	-0.9%	$0.9 ($3.0)	0.3%
2021 (Jan–Dec)	-$3.9** ($2.0)	-1.3%	-$0.9 ($2.3)	-0.3%	-$3.0 ($3.0)	-1.0%	$0.2 ($2.1)	0.1%	$3.1 ($2.4)	0.9%	-$2.8 ($3.2)	-0.9%
2020–2021 (combined COVID years)	-$3.8** ($1.7)	-1.2%	-$2.8 ($1.9)	-0.9%	-$1.0 ($2.5)	-0.3%	-$0.9 ($1.8)	-0.3%	$0.0 ($2.0)	-0.01%	-$0.9 ($2.7)	-0.3%
Outpatient ED visits, including observation stays												
2019 (Jan–Dec)	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA
2020 (Jan–Dec)	-0.5 (2.3)	-0.1%	2.0 (2.3)	0.4%	-2.5 (3.2)	-0.5%	-1.6 (2.1)	-0.3%	-5.1** (2.3)	-1.0%	3.5 (3.1)	0.7%
2021 (Jan–Dec)	-1.9 (2.6)	-0.4%	5.9** (2.6)	1.1%	-7.8** (3.7)	-1.6%	-3.4 (2.4)	-0.7%	-4.1 (2.8)	-0.8%	0.6 (3.7)	0.1%
2020–2021 (combined COVID years)	-1.2 (2.3)	-0.2%	3.9* (2.2)	0.7%	-5.1 (3.2)	-1.0%	-2.5 (2.1)	-0.5%	-4.6** (2.3)	-0.9%	2.1 (3.1)	0.4%

Table 5.D.vii. ***(continued)***

	Track 1, CPC+ vs. comparison differences relative to 2019		Track 1, Non-CPC+ vs. non-comparison differences relative to 2019		Track 1, Triple-differences model		Track 2, CPC+ vs. comparison differences relative to 2019		Track 2, Non-CPC+ vs. non-comparison differences relative to 2019		Track 2, Triple-differences model	
	Estimate (SE)	Percentage difference relative to 2019 comparison mean[a]	Estimate (SE)	Percentage difference relative to 2019 non-comparison mean[b]	Estimate (SE)	Percentage difference relative to 2019 comparison mean[a]	Estimate (SE)	Percentage difference relative to 2019 non-comparison mean[b]	Estimate (SE)	Percentage difference relative to 2019 comparison mean[a]	Estimate (SE)	Percentage difference relative to 2019 non-comparison mean[b]
Expenditures on outpatient ED visits, including observation stays (per beneficiary per month)												
2019 (Jan–Dec)	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA
2020 (Jan–Dec)	-$0.1 ($0.2)	-0.3%	-$0.1 ($0.2)	-0.3%	$0.0 ($0.3)	-0.1%	-$0.2 ($0.2)	-0.8%	-$0.5*** ($0.2)	-1.6%	$0.3 ($0.3)	0.9%
2021 (Jan–Dec)	-$0.2 ($0.3)	-0.7%	-$0.2 ($0.3)	-0.6%	$0.0 ($0.4)	-0.1%	-$0.5** ($0.3)	-1.7%	-$0.5* ($0.3)	-1.4%	$0.0 ($0.4)	-0.1%
2020–2021 (combined COVID years)	-$0.2 ($0.2)	-0.5%	-$0.1 ($0.2)	-0.4%	$0.0 ($0.3)	-0.1%	-$0.4* ($0.2)	-1.2%	-$0.5** ($0.2)	-1.5%	$0.1 ($0.3)	0.4%
Urgent care center visits (per 1,000 beneficiaries per year)												
2019 (Jan–Dec)	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA
2020 (Jan–Dec)	14.5*** (3.0)	9.5%	13.4*** (1.6)	10.4%	1.1 (3.4)	0.7%	7.9*** (2.8)	5.5%	8.1*** (1.8)	6.4%	-0.2 (3.3)	-0.1%
2021 (Jan–Dec)	15.4*** (4.4)	10.1%	20.2*** (2.6)	15.6%	-4.8 (5.1)	-3.2%	11.4** (4.7)	7.9%	15.8*** (2.7)	12.5%	-4.4 (5.4)	-3.0%
2020–2021 (combined COVID years)	14.9*** (3.4)	9.8%	16.7*** (2.0)	12.9%	-1.8 (4.0)	-1.2%	9.7*** (3.5)	6.7%	11.9*** (2.1)	9.4%	-2.2 (4.0)	-1.5%

Source: Mathematica's analysis of Medicare claims data from January 2019 through December 2021.

Note: Estimates in the table are derived from models run at the practice-month-year level that are regression-adjusted for baseline beneficiary characteristics and practice fixed effects. For CPC+ practices, observations were weighted by the number of Medicare FFS beneficiaries assigned to the practice during the month and year. For comparison practices, the weight is a product of the number of assigned beneficiaries and the matching weight. For non-CPC+ and non-comparison practices, we used a concentration weight constructed at the state-HRR level such that non-CPC+ practices had the same level of representation (in terms of beneficiary months) as CPC+ practices in the same state and HRR and SSP group, and non-comparison practices had the same level of representation as comparison practices in the same state and HRR and SSP group. Among non-CPC+ and non-comparison practices, we winsorized the weights at the 99th percentile. Standard errors are clustered at the practice level. For a detailed description of methods, see Supplement 5.

*/**/*** Significantly different from zero at the 0.10/0.05/0.01 levels, two-tailed test.

[a] To calculate these percentages, we divided the estimated CPC+ versus comparison differences by the unadjusted 2019 comparison mean for the outcome.

[b] To calculate these percentages, we divided the estimated non-CPC+ versus non-comparison differences by the unadjusted 2019 non-comparison mean for the outcome.

CPC+ = Comprehensive Primary Care Plus; ED = emergency department; FFS = fee-for-service; HCC = hierarchical condition category; HRR = hospital referral region; NA = not applicable; SE = standard error; SSP = Medicare Shared Savings Program.

Table 5.D.viii. Triple-differences models generally increased the magnitude of the differences between CPC+ and comparison practices for SSP practices and reduced differences for non-SSP practices (Track 1)

	Track 1 SSP, CPC+ vs. comparison differences relative to 2019		Track 1 SSP, Non-CPC+ vs. non-comparison differences relative to 2019		Track 1 SSP, Triple-differences model		Track 1 Non-SSP, CPC+ vs. comparison differences relative to 2019		Track 1 Non-SSP, Non-CPC+ vs. non-comparison differences relative to 2019		Track 1 Non-SSP, Triple-differences model	
	Estimate (SE)	Percentage difference relative to 2019 comparison mean[a]	Estimate (SE)	Percentage difference relative to 2019 non-comparison mean[b]	Estimate (SE)	Percentage difference relative to 2019 comparison mean[a]	Estimate (SE)	Percentage difference relative to 2019 comparison mean[a]	Estimate (SE)	Percentage difference relative to 2019 non-comparison mean[b]	Estimate (SE)	Percentage difference relative to 2019 comparison mean[a]
Medicare Part A and B expenditures without enhanced payments for CPC+ and SSP												
2019 (Jan–Dec)	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA
2020 (Jan–Dec)	-$2.4 ($4.9)	-0.2%	-$5.8 ($5.9)	-0.5%	$3.5 ($7.6)	0.3%	-$12.8*** ($4.9)	-1.3%	-$15.2*** ($4.5)	-1.5%	$2.4 ($6.6)	0.3%
2021 (Jan–Dec)	-$4.2 ($5.3)	-0.4%	$4.6 ($7.0)	0.4%	-$8.8 ($8.8)	-0.9%	-$16.6*** ($5.7)	-1.7%	-$19.9*** ($5.1)	-1.9%	$3.2 ($7.7)	0.3%
2020–2021 (combined COVID years)	-$3.3 ($4.5)	-0.3%	-$0.7 ($5.6)	-0.1%	-$2.5 ($7.2)	-0.3%	-$14.7*** ($4.7)	-1.5%	-$17.5*** ($4.2)	-1.7%	$2.8 ($6.3)	0.3%
Acute hospitalizations (short-stay acute care and critical access hospitals)												
2019 (Jan–Dec)	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA
2020 (Jan–Dec)	0.9 (1.7)	0.3%	-2.4 (2.3)	-0.8%	3.3 (2.9)	1.2%	-2.7 (1.9)	-1.0%	-3.5** (1.7)	-1.2%	0.8 (2.5)	0.3%
2021 (Jan–Dec)	0.5 (1.9)	0.2%	4.0 (2.6)	1.3%	-3.5 (3.3)	-1.2%	-2.6 (2.1)	-0.9%	-2.1 (1.9)	-0.7%	-0.5 (2.8)	-0.2%
2020–2021 (combined COVID years)	0.7 (1.6)	0.2%	0.7 (2.2)	0.2%	0.0 (2.7)	0.00%	-2.7 (1.8)	-1.0%	-2.8* (1.6)	-0.9%	0.2 (2.4)	0.1%
Expenditures on acute hospitalizations (short-stay acute care and critical access hospitals) (per beneficiary per month)												
2019 (Jan–Dec)	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA
2020 (Jan–Dec)	-$0.5 ($2.7)	-0.2%	-$4.1 ($3.4)	-1.2%	$3.6 ($4.3)	1.2%	-$6.0** ($2.6)	-2.0%	-$5.2* ($2.7)	-1.6%	-$0.8 ($3.7)	-0.3%
2021 (Jan–Dec)	-$0.7 ($2.8)	-0.2%	$3.8 ($3.6)	1.1%	-$4.5 ($4.6)	-1.5%	-$6.5** ($2.9)	-2.2%	-$6.0** ($2.7)	-1.9%	-$0.5 ($4.0)	-0.2%
2020–2021 (combined COVID years)	-$0.6 ($2.4)	-0.2%	-$0.2 ($3.0)	-0.1%	-$0.4 ($3.8)	-0.1%	-$6.3*** ($2.4)	-2.1%	-$5.6** ($2.3)	-1.7%	-$0.7 ($3.3)	-0.2%
Outpatient ED visits, including observation stays												
2019 (Jan–Dec)	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA
2020 (Jan–Dec)	-2.0 (3.0)	-0.4%	-0.3 (4.1)	-0.1%	-1.7 (5.1)	-0.4%	1.7 (3.4)	0.3%	4.1 (2.7)	0.7%	-2.4 (4.3)	-0.5%
2021 (Jan–Dec)	-1.5 (3.5)	-0.3%	6.7* (4.0)	1.3%	-8.2 (5.3)	-1.7%	-1.9 (3.9)	-0.4%	4.3 (3.2)	0.8%	-6.1 (5.0)	-1.2%
2020–2021 (combined COVID years)	-1.7 (3.0)	-0.4%	3.1 (3.6)	0.6%	-4.8 (4.7)	-1.0%	-0.1 (3.4)	-0.02%	4.2 (2.7)	0.8%	-4.3 (4.3)	-0.8%

Table 5.D.viii. ***(continued)***

	Track 1 SSP, CPC+ vs. comparison differences relative to 2019		Track 1 SSP, Non-CPC+ vs. non-comparison differences relative to 2019		Track 1 SSP, Triple-differences model		Track 1 Non-SSP, CPC+ vs. comparison differences relative to 2019		Track 1 Non-SSP, Non-CPC+ vs. non-comparison differences relative to 2019		Track 1 Non-SSP, Triple-differences model	
	Estimate (SE)	Percentage difference relative to 2019 comparison mean[a]	Estimate (SE)	Percentage difference relative to 2019 non-comparison mean[b]	Estimate (SE)	Percentage difference relative to 2019 comparison mean[a]	Estimate (SE)	Percentage difference relative to 2019 comparison mean[a]	Estimate (SE)	Percentage difference relative to 2019 non-comparison mean[b]	Estimate (SE)	Percentage difference relative to 2019 comparison mean[a]
Expenditures on outpatient ED visits, including observation stays (per beneficiary per month)												
2019 (Jan–Dec)	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA
2020 (Jan–Dec)	-$0.1 ($0.3)	-0.2%	-$0.4 ($0.3)	-1.2%	$0.3 ($0.4)	1.1%	-$0.2 ($0.3)	-0.5%	$0.2 ($0.2)	0.6%	-$0.4 ($0.4)	-1.2%
2021 (Jan–Dec)	-$0.2 ($0.3)	-0.7%	-$1.0** ($0.4)	-3.2%	$0.8 ($0.5)	2.7%	-$0.2 ($0.4)	-0.7%	$0.5 ($0.4)	1.4%	-$0.7 ($0.6)	-2.3%
2020–2021 (combined COVID years)	-$0.1 ($0.3)	-0.5%	-$0.7** ($0.3)	-2.2%	$0.5 ($0.4)	1.9%	-$0.2 ($0.3)	-0.6%	$0.3 ($0.3)	1.0%	-$0.5 ($0.4)	-1.7%
Urgent care center visits (per 1,000 beneficiaries per year)												
2019 (Jan–Dec)	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA
2020 (Jan–Dec)	18.0*** (3.9)	11.3%	18.1*** (2.7)	12.9%	-0.1 (4.7)	-0.1%	10.9** (4.6)	7.6%	9.8*** (1.8)	8.3%	1.1 (4.9)	0.8%
2021 (Jan–Dec)	23.5*** (6.3)	14.8%	26.1*** (4.5)	18.6%	-2.6 (7.7)	-1.6%	8.1 (5.9)	5.6%	14.5*** (2.8)	12.2%	-6.4 (6.5)	-4.4%
2020–2021 (combined COVID years)	20.7*** (4.7)	13.1%	22.0*** (3.3)	15.7%	-1.3 (5.7)	-0.8%	9.5* (5.0)	6.6%	12.1*** (2.1)	10.2%	-2.6 (5.4)	-1.8%

Source: Mathematica's analysis of Medicare claims data from January 2019 through December 2021.

Note: Estimates in the table are derived from models run at the practice-month-year level that are regression-adjusted for baseline beneficiary characteristics and practice fixed effects. For CPC+ practices, observations were weighted by the number of Medicare FFS beneficiaries assigned to the practice during the month and year. For comparison practices, the weight is a product of the number of assigned beneficiaries and the matching weight. For non-CPC+ and non-comparison practices, we used a concentration weight constructed at the state-HRR level such that non-CPC+ practices had the same level of representation (in terms of beneficiary months) as CPC+ practices in the same state and HRR and SSP group, and non-comparison practices had the same level of representation as comparison practices in the same state and HRR and SSP group. Among non-CPC+ and non-comparison practices, we winsorized the weights at the 99th percentile. Standard errors are clustered at the practice level. For a detailed description of methods, see Supplement 5.

*/**/*** Significantly different from zero at the 0.10/0.05/0.01 levels, two-tailed test.

[a] To calculate these percentages, we divided the estimated CPC+ versus comparison differences by the unadjusted 2019 comparison mean for the outcome.

[b] To calculate these percentages, we divided the estimated non-CPC+ versus non-comparison differences by the unadjusted 2019 non-comparison mean for the outcome.

CPC+ = Comprehensive Primary Care Plus; ED = emergency department; FFS = fee-for-service; HCC = hierarchical condition category; HRR = hospital referral region; NA = not applicable; SE = standard error; SSP = Medicare Shared Savings Program.

Table 5.D.ix. Triple-differences models generally reduced the magnitude of the differences between CPC+ and comparison practices for SSP practices and increased differences for non-SSP practices (Track 2)

	Track 1 SSP, CPC+ vs. comparison differences relative to 2019		Track 1 SSP, Non-CPC+ vs. non-comparison differences relative to 2019		Track 1 SSP, Triple-differences model		Track 1 Non-SSP, CPC+ vs. comparison differences relative to 2019		Track 1 Non-SSP, Non-CPC+ vs. non-comparison differences relative to 2019		Track 1 Non-SSP, Triple-differences model	
	Estimate (SE)	Percentage difference relative to 2019 comparison mean[a]	Estimate (SE)	Percentage difference relative to 2019 non-comparison mean[b]	Estimate (SE)	Percentage difference relative to 2019 comparison mean[a]	Estimate (SE)	Percentage difference relative to 2019 comparison mean[a]	Estimate (SE)	Percentage difference relative to 2019 non-comparison mean[b]	Estimate (SE)	Percentage difference relative to 2019 comparison mean[a]
Medicare Part A and B expenditures without enhanced payments for CPC+ and SSP												
2019 (Jan–Dec)	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA
2020 (Jan–Dec)	-$14.2*** ($4.9)	-1.4%	-$11.4* ($6.4)	-1.1%	-$2.8 ($8.1)	-0.3%	-$3.3 ($4.6)	-0.3%	-$7.3* ($4.4)	-0.7%	$4.0 ($6.4)	0.4%
2021 (Jan–Dec)	-$13.9** ($5.7)	-1.4%	-$4.9 ($7.7)	-0.5%	-$9.0 ($9.6)	-0.9%	-$2.4 ($5.5)	-0.3%	-$13.2*** ($5.1)	-1.3%	$10.8 ($7.5)	1.1%
2020–2021 (combined COVID years)	-$14.1*** ($4.7)	-1.4%	-$8.2 ($6.1)	-0.8%	-$5.8 ($7.7)	-0.6%	-$2.9 ($4.5)	-0.3%	-$10.2** ($4.2)	-1.0%	$7.3 ($6.1)	0.8%
Acute hospitalizations (short-stay acute care and critical access hospitals)												
2019 (Jan–Dec)	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA
2020 (Jan–Dec)	-4.4** (2.0)	-1.5%	-2.9 (2.6)	-1.0%	-1.5 (3.3)	-0.5%	0.7 (1.7)	0.2%	-0.7 (1.6)	-0.2%	1.4 (2.3)	0.5%
2021 (Jan–Dec)	-1.7 (2.1)	-0.6%	5.4* (3.0)	1.7%	-7.1* (3.7)	-2.5%	1.4 (1.9)	0.5%	-0.4 (1.8)	-0.1%	1.8 (2.6)	0.6%
2020–2021 (combined COVID years)	-3.1* (1.9)	-1.1%	1.1 (2.5)	0.4%	-4.2 (3.1)	-1.4%	1.0 (1.6)	0.4%	-0.6 (1.5)	-0.2%	1.6 (2.2)	0.6%
Expenditures on acute hospitalizations (short-stay acute care and critical access hospitals) (per beneficiary per month)												
2019 (Jan–Dec)	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA
2020 (Jan–Dec)	-$3.6 ($3.0)	-1.1%	-$3.0 ($3.6)	-0.9%	-$0.6 ($4.7)	-0.2%	-$1.5 ($2.7)	-0.5%	-$2.8 ($2.7)	-0.9%	$1.3 ($3.8)	0.4%
2021 (Jan–Dec)	-$0.8 ($3.1)	-0.2%	$9.3** ($4.0)	2.7%	-$10.0** ($5.1)	-3.2%	-$0.4 ($3.0)	-0.1%	-$2.6 ($2.9)	-0.8%	$2.2 ($4.2)	0.7%
2020–2021 (combined COVID years)	-$2.2 ($2.6)	-0.7%	$3.0 ($3.3)	0.9%	-$5.2 ($4.2)	-1.6%	-$0.9 ($2.5)	-0.3%	-$2.7 ($2.4)	-0.8%	$1.8 ($3.5)	0.6%
Outpatient ED visits, including observation stays												
2019 (Jan–Dec)	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA
2020 (Jan–Dec)	-10.2*** (3.0)	-2.2%	-11.4*** (4.4)	-2.2%	1.2 (5.3)	0.3%	6.2** (2.8)	1.2%	-1.9 (2.7)	-0.4%	8.2** (3.9)	1.6%
2021 (Jan–Dec)	-14.1*** (3.4)	-3.0%	-9.4* (5.0)	-1.8%	-4.7 (6.1)	-1.0%	6.4* (3.4)	1.3%	-1.8 (3.3)	-0.3%	8.2* (4.7)	1.6%
2020–2021 (combined COVID years)	-12.2*** (3.0)	-2.6%	-10.4** (4.2)	-2.0%	-1.8 (5.1)	-0.4%	6.3** (2.8)	1.3%	-1.9 (2.7)	-0.3%	8.2** (3.9)	1.6%

Table 5.D.ix. ***(continued)***

	Track 1 SSP, CPC+ vs. comparison differences relative to 2019		Track 1 SSP, Non-CPC+ vs. non-comparison differences relative to 2019		Track 1 SSP, Triple-differences model		Track 1 Non-SSP, CPC+ vs. comparison differences relative to 2019		Track 1 Non-SSP, Non-CPC+ vs. non-comparison differences relative to 2019		Track 1 Non-SSP, Triple-differences model	
	Estimate (SE)	Percentage difference relative to 2019 comparison mean[a]	Estimate (SE)	Percentage difference relative to 2019 non-comparison mean[b]	Estimate (SE)	Percentage difference relative to 2019 comparison mean[a]	Estimate (SE)	Percentage difference relative to 2019 comparison mean[a]	Estimate (SE)	Percentage difference relative to 2019 non-comparison mean[b]	Estimate (SE)	Percentage difference relative to 2019 comparison mean[a]
Expenditures on outpatient ED visits, including observation stays (per beneficiary per month)												
2019 (Jan–Dec)	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA
2020 (Jan–Dec)	-$0.3 ($0.3)	-1.1%	-$1.3*** ($0.4)	-3.8%	$0.9** ($0.5)	3.0%	-$0.2 ($0.3)	-0.5%	-$0.1 ($0.2)	-0.3%	-$0.1 ($0.3)	-0.2%
2021 (Jan–Dec)	-$1.2*** ($0.4)	-3.9%	-$1.9*** ($0.5)	-5.9%	$0.7 ($0.6)	2.4%	$0.1 ($0.3)	0.4%	$0.3 ($0.3)	0.8%	-$0.1 ($0.5)	-0.5%
2020–2021 (combined COVID years)	-$0.8** ($0.3)	-2.5%	-$1.6*** ($0.4)	-4.8%	$0.8* ($0.5)	2.7%	$0.0 ($0.3)	-0.03%	$0.1 ($0.3)	0.3%	-$0.1 ($0.4)	-0.3%
Urgent care center visits (per 1,000 beneficiaries per year)												
2019 (Jan–Dec)	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA
2020 (Jan–Dec)	7.1* (4.1)	4.9%	7.5** (3.4)	5.5%	-0.4 (5.3)	-0.3%	8.3** (3.8)	5.7%	8.4*** (2.0)	7.1%	-0.1 (4.3)	-0.1%
2021 (Jan–Dec)	5.6 (7.5)	3.8%	13.3*** (5.1)	9.8%	-7.7 (9.1)	-5.3%	15.9*** (6.0)	11.0%	18.0*** (3.1)	15.1%	-2.0 (6.7)	-1.4%
2020–2021 (combined COVID years)	6.3 (5.4)	4.4%	10.3*** (3.9)	7.6%	-4.0 (6.7)	-2.7%	12.1*** (4.6)	8.3%	13.1*** (2.3)	11.0%	-1.0 (5.2)	-0.7%

Source: Mathematica's analysis of Medicare claims data from January 2019 through December 2021.

Note: Estimates in the table are derived from models run at the practice-month-year level that are regression-adjusted for baseline beneficiary characteristics and practice fixed effects. For CPC+ practices, observations were weighted by the number of Medicare FFS beneficiaries assigned to the practice during the month and year. For comparison practices, the weight is a product of the number of assigned beneficiaries and the matching weight. For non-CPC+ and non-comparison practices, we used a concentration weight constructed at the state-HRR level such that non-CPC+ practices had the same level of representation (in terms of beneficiary months) as CPC+ practices in the same state and HRR and SSP group, and non-comparison practices had the same level of representation as comparison practices in the same state and HRR and SSP group. Among non-CPC+ and non-comparison practices, we winsorized the weights at the 99th percentile. Standard errors are clustered at the practice level.

*/**/*** Significantly different from zero at the 0.10/0.05/0.01 levels, two-tailed test.

[a] To calculate these percentages, we divided the estimated CPC+ versus comparison differences by the unadjusted 2019 comparison mean for the outcome.

[b] To calculate these percentages, we divided the estimated non-CPC+ versus non-comparison differences by the unadjusted 2019 non-comparison mean for the outcome.

CPC+ = Comprehensive Primary Care Plus; ED = emergency department; FFS = fee-for-service; HCC = hierarchical condition category; HRR = hospital referral region; NA = not applicable; SE = standard error; SSP = Medicare Shared Savings Program.

Supplement 8. COVID-19 and the analysis sample composition

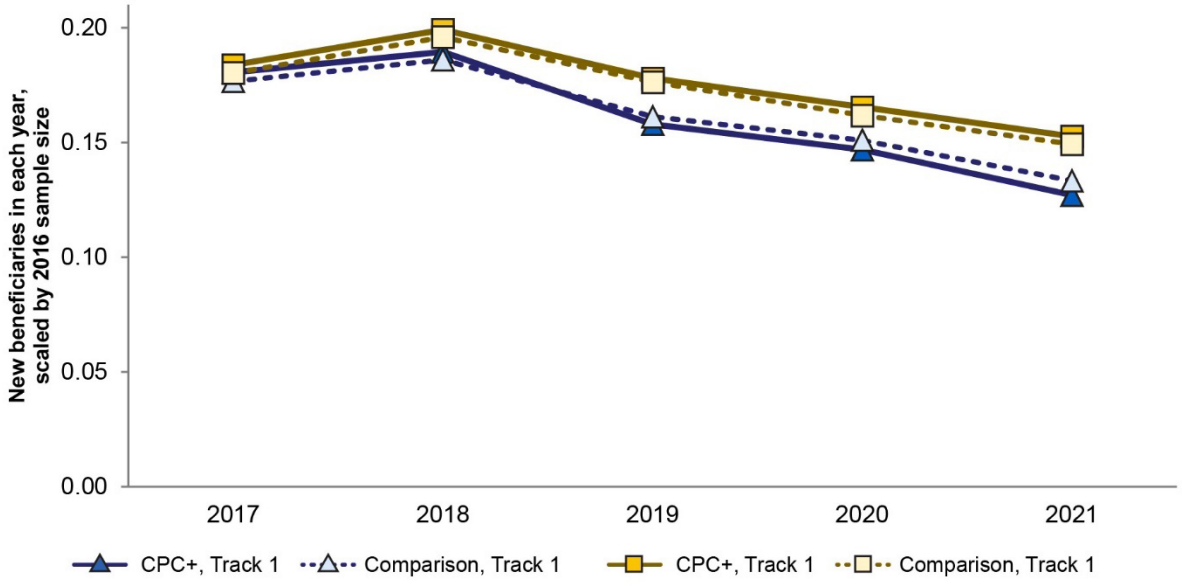
In this section, we provide additional information about our analysis sample in relation to the claims-based COVID-19 checks. In Section A, we examine whether there were differential changes in the composition of the beneficiary sample between CPC+ and comparison practices during the pandemic period. Similarly, Section B checks whether the sample of practices changed differentially between CPC+ and comparison groups, for example, due to differential rates of practice closure during the pandemic.

A. Beneficiary counts and characteristics

We examined changes in the number of attributed beneficiaries for CPC+ and comparison practices over time because it is possible that these numbers could differentially change during the COVID-19 pandemic due to difference in attribution-eligible visits. We found that beneficiary attribution changed in a similar way between CPC+ and comparison practices over time, by track, and within each track, by SSP status.

Specifically, we examined the percentages of beneficiaries added to the sample each year (Figures 5.D.xi and 5.D.xii). There was a relatively large decline between 2018 and 2019, due to the changes in practice composition (in terms of practitioners) driven by the switch from SK&A to OneKey data in 2019. The subsequent decreases between 2019 and 2020 and between 2020 and 2021 were relatively small and similar in magnitude for CPC+ and comparison groups, overall and by SSP status. The small decrease in the percentage of beneficiaries added in 2020 and 2021 may be because there were fewer primary care visits those years (which in turn leads to fewer beneficiaries newly assigned to practices) during the pandemic.

Figure 5.D.xi. Fewer beneficiaries were added to the sample in 2020 and 2021, but the changes were similar between CPC+ and comparison practices by track

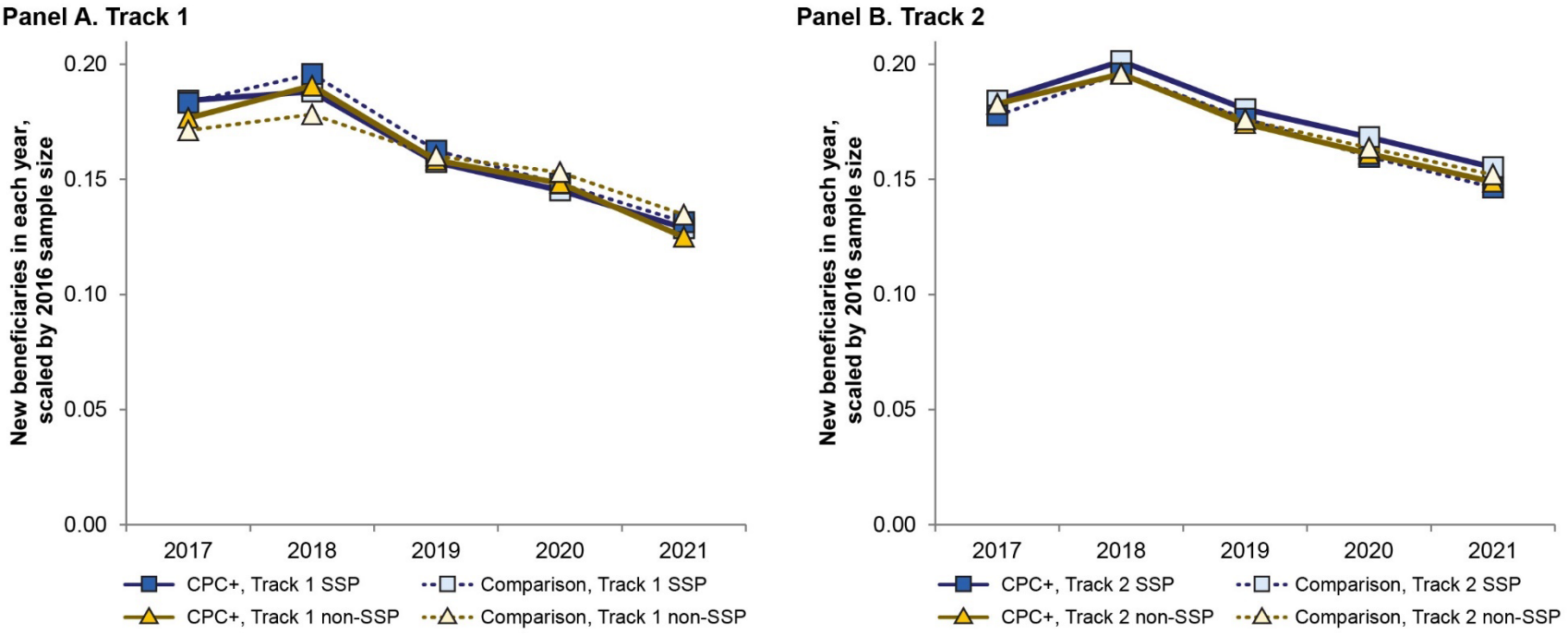


Source: Mathematica's analysis of Medicare claims data from January 2016 through December 2021.

Note: The percentages are calculated among the total number of beneficiaries included in the analytic sample each year. Beneficiaries could be added to the sample if they switched from health maintenance organizations, newly enrolled in Medicare Part A and B, newly had Medicare as their primary payer, or were newly assigned to CPC+ or comparison practices in the reference year.

CPC+ = Comprehensive Primary Care Plus.

Figure 5.D.xii. Fewer beneficiaries were added to the sample in 2020 and 2021, but the changes were similar between CPC+ and comparison practices by SSP status



Source: Mathematica's analysis of Medicare claims data from January 2016 through December 2021.

Note: The percentages are calculated among the total number of beneficiaries included in the analytic sample each year. Beneficiaries could be added to the sample if they switched from health maintenance organizations, newly enrolled in Medicare Part A and B, newly had Medicare as their primary payer, or were newly assigned to CPC+ or comparison practices in the reference year.

CPC+ = Comprehensive Primary Care Plus; SSP = Medicare Shared Savings Program.

We also studied the change in characteristics of Medicare beneficiaries in CPC+ and comparison practices over time (Table 5.D.x). COVID-19 might have affected beneficiary characteristics for multiple reasons such as sicker patients dying (so no longer fee-for-service [FFS] eligible for the sample), altering individuals' decision to enroll in FFS versus managed care, changing the types of newly attributed patients joining the sample, or affecting the health status of the beneficiaries. To study whether this occurred, we allowed characteristics to be updated each year for a given beneficiary (for example, their hierarchical condition category score). Therefore, any changes in observed mean characteristics from one year to the next could reflect changes in beneficiaries who are in the sample or changes in the characteristics of beneficiaries who remain in the sample. We did not find evidence COVID-19 altered the characteristics of beneficiaries in the ITT sample, and the difference between CPC+ and comparison practices were small (less than 0.02 standardized difference) for all characteristics in all years (Table 5.D.x).

Table 5.D.x. Beneficiary characteristics (which are updated each year) were similar for CPC+ and comparison practices (characteristics shown as percentage, unless otherwise noted)

	2019				2020				2021			
Measure	CPC+[a]	Comparison[a]	Difference	Standardized difference	CPC+[a]	Comparison[a]	Difference	Standardized difference	CPC+[a]	Comparison[a]	Difference	Standardized difference
Track 1												
Age (mean)	72.1	72.2	0.0	0.00	72.4	72.5	-0.1	-0.01	72.7	72.8	-0.1	-0.01
Race												
White	88.1	87.8	0.3	0.01	88.1	87.8	0.4	0.01	88.2	87.8	0.4	0.01
Black	5.2	5.5	-0.4	-0.02	4.9	5.3	-0.4	-0.02	4.6	5.1	-0.5	-0.02
Other	6.7	6.7	0.0	0.00	7.0	6.9	0.1	0.00	7.2	7.1	0.0	0.00
Male	42.2	42.3	-0.1	0.00	42.3	42.4	-0.1	0.00	42.3	42.4	-0.1	0.00
Original reason for Medicare eligibility												
Age	80.5	80.6	0.0	0.00	81.2	81.3	-0.1	0.00	82.1	82.1	0.0	0.00
Disabled	19.0	18.9	0.1	0.00	18.3	18.2	0.1	0.00	17.4	17.4	0.0	0.00
ESRD	0.5	0.5	0.0	-0.01	0.5	0.6	-0.1	-0.01	0.5	0.5	0.0	-0.01
Chronic conditions												
Diabetes with chronic complications	15.6	15.6	0.0	0.00	16.1	16.2	-0.1	0.00	15.6	15.7	-0.2	-0.01
Chronic obstructive pulmonary disease	12.4	11.8	0.6	0.02	12.5	11.7	0.7	0.02	11.3	10.7	0.6	0.02
Congestive heart failure	11.8	11.8	0.0	0.00	12.3	12.3	0.0	0.00	11.8	11.7	0.1	0.00
Ischemic heart disease, acute myocardial infarction, angina	6.1	5.6	0.5	0.02	6.4	5.9	0.5	0.02	5.9	5.5	0.4	0.02
HCC score[b]	1.2	1.2	0.0	0.01	1.2	1.2	0.0	0.01	1.1	1.1	0.0	0.00
Track 2												
Age (mean)	72.0	72.0	0.0	0.00	72.3	72.3	0.0	0.00	72.6	72.6	0.0	0.00
Race												
White	87.2	87.2	0.0	0.00	87.2	87.3	0.0	0.00	87.3	87.3	0.0	0.00
Black	5.8	6.1	-0.3	-0.01	5.6	5.9	-0.3	-0.01	5.3	5.6	-0.3	-0.01
Other	7.0	6.7	0.3	0.01	7.2	6.9	0.3	0.01	7.4	7.1	0.3	0.01
Male	42.4	42.4	0.0	0.00	42.4	42.5	-0.1	0.00	42.5	42.6	-0.1	0.00
Original reason for Medicare eligibility												
Age	80.8	80.4	0.4	0.01	81.5	81.1	0.4	0.01	82.4	82.0	0.4	0.01
Disabled	18.7	19.0	-0.3	-0.01	18.0	18.3	-0.3	-0.01	17.1	17.5	-0.4	-0.01
ESRD	0.5	0.6	0.0	-0.01	0.5	0.6	-0.1	-0.01	0.5	0.5	-0.1	-0.01

Table 5.D.x. ***(continued)***

	2019				2020				2021			
Measure	CPC+[a]	Comparison[a]	Difference	Standardized difference	CPC+[a]	Comparison[a]	Difference	Standardized difference	CPC+[a]	Comparison[a]	Difference	Standardized difference
Chronic conditions												
Diabetes with chronic complications	16.1	15.5	0.6	0.02	16.7	16.2	0.5	0.01	16.2	15.7	0.4	0.01
Chronic obstructive pulmonary disease	12.0	11.5	0.6	0.02	12.1	11.5	0.6	0.02	11.0	10.4	0.5	0.02
Congestive heart failure	12.0	11.7	0.3	0.01	12.5	12.2	0.3	0.01	11.9	11.7	0.2	0.01
Ischemic heart disease, acute myocardial infarction, angina	6.0	5.5	0.5	0.02	6.3	5.9	0.4	0.02	5.9	5.5	0.3	0.01
HCC score[b]	1.2	1.2	0.02	0.01	1.2	1.2	0.0	0.01	1.2	1.1	0.01	0.01

Source: Mathematica's analysis of Medicare claims and enrollment data for January 2019 through December 2021.

Note: Characteristics were measured as of January 1 of each calendar year, using a two-year lookback period for chronic conditions.
Sample sizes. In Track 1 for CPC+ and comparison practices, respectively, this analysis includes 1,041,306 and 3,538,453 beneficiaries in 2019; 1,070,852 and 3,708,366 beneficiaries in 2020; and 1,072,211 and 3,774,922 beneficiaries in 2021. In Track 2 for CPC+ and comparison practices, respectively, this analysis includes 1,267,844 and 2,996,248 beneficiaries in 2019; 1,312,712 and 3,133,710 beneficiaries in 2020; and 1,323,655 and 3,184,540 beneficiaries in 2021.

[a] Means were weighted to account for (1) the share of the year for which the beneficiary's data were observed and (2) the matching (for beneficiaries in comparison practices only).

[b] HCC scores are a measure of risk for subsequent expenditures. The Centers for Medicare & Medicaid Services calculates them such that the average for the Medicare fee-for-service population nationally is 1.0. A patient with a risk score of 1.30 is predicted to have expenditures that would be approximately 30 percent above the average, whereas a patient with a risk score of 0.70 is expected to have expenditures that would be approximately 30 percent below the average.

CPC+ = Comprehensive Primary Care Plus; ESRD = end-stage renal disease; HCC = hierarchical condition category.

B. Practice counts and service interruptions

The COVID-19 pandemic could have affected the number and composition of practices included in the ITT sample, for example, due to practice closures or practitioner turnover driven by financial or operational reasons. CPC+ payment supports could have also influenced practices' ability to continue operating or their ability to retain certain staff, particularly during the pandemic. Therefore, to understand potential effects of COVID-19 on practice closure and practice compositional changes during 2020, we examined trends in the number of practices and practitioners for both the CPC+ and comparison groups over time, using IQVIA data on practice composition.

Over the six-year period, trends in the number of practices (a general decline) and practice size (an increase in the number of primary care practitioners) were similar for CPC+ and comparison practices overall and when stratifying by SSP status (Table 5.D.xi). Changes between 2018 and 2019 are likely explained by the change in IQVIA's source data from SK&A (2016–2018) to OneKey (2019–2021). We did not see any substantial differences in the number or proportion of CPC+ and comparison practices exiting the sample between 2019 and 2020 or between 2019 and 2021 relative to historical trends, which suggests that COVID-19 likely did not have a major effect on practice closures or mergers in our evaluation sample. For example, 21 practices (0.8 percent) fell out of the CPC+ sample between 2019 and 2020, and 44 practices (1.6 percent) fell out between 2019 and 2021 (including the 21 practices that dropped in 2020). Before the pandemic, 13 practices (0.5 percent) fell out of the CPC+ sample between 2016 and 2017, and 24 practices (0.8 percent) fell out between 2016 and 2018. Compared to CPC+ practices, comparison practices fell out at higher rates before and during the pandemic (2.9 percent of comparison practices fell out between 2016 and 2018 and 3.2 percent between 2019 and 2021).

Table 5.D.xi. The numbers of practices and primary care practitioners in 2020 and 2021 were not substantially different from historical trends for CPC+ and comparison practices, with large changes from 2018 to 2019 likely explained by changes to the data source

	2016	2017	2018	2019	2020	2021	Percentage change from 2016 to 2018	Percentage change from 2019 to 2021
All practices (SSP and non-SSP combined)								
CPC+ practices								
Number of practices	2,888	2,875	2,864	2,703	2,682	2,659	-0.8	-1.6
Primary care practitioners	12,404	12,970	13,421	16,850	17,180	18,035	8.2	7.0
PCP per practice	4.3	4.5	4.7	6.2	6.4	6.8	9.1	8.8
Comparison practices								
Number of practices	6,921	6,782	6,723	5,999	5,908	5,807	-2.9	-3.2
Primary care practitioners	28,302	28,673	29,437	36,668	38,057	39,069	4.0	6.5
PCP per practice	4.1	4.2	4.4	6.1	6.4	6.7	7.1	10.1
SSP practices								
CPC+ practices								
Number of practices	1,374	1,367	1,359	1273	1,256	1,244	-1.1	-2.3
Primary care practitioners	5,747	6,140	6,276	7,701	8,060	8,422	9.2	9.4
PCP per practice	4.2	4.5	4.6	6.1	6.4	6.8	10.4	11.9
Comparison practices								
Number of practices	3,635	3,560	3,529	3,181	3,150	3,091	-2.9	-2.8
Primary care practitioners	14,896	15,107	15,582	19,257	19,802	20,576	4.6	6.8
PCP per practice	4.1	4.2	4.4	6.0	6.3	6.7	7.7	10.0
Non-SSP practices								
CPC+ practices								
Number of practices	1,514	1,508	1,505	1,430	1,426	1,415	-0.6	-1.0
Primary care practitioners	6,657	6,830	7,145	9,149	9,120	9,613	7.3	5.1
PCP per practice	4.4	4.5	4.7	6.4	6.4	6.8	8.0	6.2
Comparison practices								
Number of practices	3,286	3,222	3,194	2,818	2,758	2,716	-2.8	-3.6
Primary care practitioners	13,406	13,566	13,855	17,411	18,255	18,493	3.3	6.2
PCP per practice	4.1	4.2	4.3	6.2	6.6	6.8	6.3	10.2

Source: 2016, 2017, and 2018 SK&A data; 2019, 2020, and 2021 OneKey data.

Note: For 148 CPC+ practices we could not identify in the 2016 SK&A data, we used information from practice rosters for all years. All statistics are unweighted. We considered practices as falling out of the sample if we did not find them in the 2017 or 2018 SK&A data, or the 2019, 2020, or 2021 OneKey data. Because of our intent-to-treat design, we did not remove CPC+ practices that withdraw from CPC+.

CPC+ = Comprehensive Primary Care Plus; PCP = primary care practitioner; SSP = Medicare Shared Savings Program.

Other analyses showed that participation in CPC+ did not have any effect during the pandemic on prolonged physician service interruptions (that is, the scenario where the physician stopped billing Medicare in a month and did not resume billing within the next six months of that year), with similar changes in interruptions for CPC+ and comparison practices (Lei et al. 2022).

5.E. Empirical strategy

This Appendix describes the empirical strategy used to estimate impacts on Medicare claims-based outcomes in this report. For the main impact analysis over the five years of CPC+, we used a difference-in-differences regression analysis with a comparison group selected using propensity score matching and reweighting methods. Our sample includes practices that joined CPC+ in 2017 and were participating in CPC+ as of April 1, 2017 (the end of the first program quarter),[63] and their matched comparison practices.

In this Appendix, we first briefly describe the approach used to select the comparison group and show the similarity between the CPC+ and matched comparison practices at baseline (Section 1). We then describe the study population and unit of observation in the regressions (Section 2). We describe the regression model, including the difference-in-differences and straight-difference models (defined below) in Section 3, and discuss the interpretation of model coefficients in Section 4. We present additional details on model estimation in Section 5, followed by a description of control variables (Section 6) and weighting (Section 7). We then discuss the power to detect effects (Section 8). Finally, we describe the subgroup analyses to check for differential effects of CPC+ on practice and beneficiary subgroups (Section 9), and sensitivity tests to check for the robustness of the impact estimates (Section 10), including the triple-differences analyses to account for regional differences in trends among CPC+ and comparison practices (Section 11).

5.E.1. Comparison group

To estimate the impact of CPC+, we compared patient outcomes over time for CPC+ practices relative to those of similar matched comparison practices. We drew the comparison group from practices that provide primary care in regions not selected for CPC+. We selected comparison groups separately for Track 1 and Track 2, because CMS views each track as a different intervention that should be analyzed separately. We also matched practices separately within track by SSP status, because we and CMS deemed participation in SSP to be the most important practice characteristic that could affect outcomes, given that SSP practices face different payment incentives. The result was six comparison groups supporting analyses for six groups: (1) Track 1 overall, (2) Track 2 overall, (3) Track 1 SSP, (4) Track 1 non-SSP, (5) Track 2 SSP, and (6) Track 2 non-SSP. Appendix 6.C in the appendix to our second annual report (Ghosh et al. 2020) contains more details on the comparison groups.

We used propensity score matching and reweighting methods to establish a group of non-participating primary care practices that had similar practice characteristics (such as the number of practitioners and urban/rural status) and that served a similar population of Medicare fee-for-service (FFS) beneficiaries at baseline as CPC+ practices (for example, in terms of average age and expenditures during the year before CPC+ began, as shown in Table 5.E.1). We identified these characteristics from Medicare claims and enrollment data as well as other secondary data sources such as IQVIA, CMS data on participation in Center for Medicare and Medicaid Innovation models other than CPC+, and the Area Health Resource File.

[63] Of the 2,905 CPC+ practices that started the initiative on January 1, 2017, 17 practices (0.6 percent) withdrew in the first three months before the selection of the intent-to-treat (ITT) sample, and 2,888 practices were participating as of April 1, 2017. These 2,888 practices are in the ITT sample; we excluded the 17 practices that withdrew in the first three months because they were unlikely to have made much progress implementing CPC+ during that time.

The resulting comparison groups had baseline characteristics comparable to the CPC+ practices, and differences between the CPC+ and comparison groups were negligible for almost all characteristics across both tracks (see Table 5.E.1). Details on the post-matching similarity of the CPC+ practices and their matched comparison practices, including standardized differences, by track and SSP status for the full set of characteristics that were used for matching are in our second annual report appendix (see Ghosh et al. 2020, Tables 6.C.5 to 6.C.10).

Table 5.E.1. Similarity of the CPC+ and comparison practices (practice values weighted by number of Medicare FFS beneficiaries), by track

		Track 1		Track 2	
Practice characteristics at baseline	Data source for characteristic	Mean among CPC+ practices (N = 1,373)	Weighted mean among comparison practices (N = 5,243)	Mean among CPC+ practices (N = 1,515)	Weighted mean among comparison practices (N = 3,783)
Participation in SSP ACO as of January 1, 2017 (%)	MDM January 1, 2017	51.4	52.3	44.2	44.2
Hospital ownership or health system management or ownership (%)	SK&A 2016	54.8	55.3	58.2	59.8
Participation in prior primary care transformation activities[a] (%)	Data from CMS and from organizations that offer medical home recognition	53.5	52.6	80.9	75.4
Urbanicity of practice's county					
Rural (%)	Area Health Resource File 2016	10.3	9.8	7.7	7.7
Suburban (%)	Area Health Resource File 2016	18.0	18.4	16.0	16.8
Urban (%)	Area Health Resource File 2016	71.7	71.8	76.3	75.5
Mean PBPM Medicare expenditures in 2016	EDB and claims data	$881.0	$885.0	$877.0	$879.0
Acute care hospitalizations (short-stay acute care and CAHs) in 2016 per 1,000 beneficiaries, annualized	EDB and claims data	285.4	284.0	287.4	283.5
Outpatient ED visits, including observation stays, in 2016 per 1,000 beneficiaries, annualized	EDB and claims data	493.8	498.2	492.6	492.5
Mean 2016 HCC score among beneficiaries assigned in 2016	EDB and claims data	1.1	1.1	1.1	1.1
Number of primary care practitioners:					
1–2 primary care practitioners (%)	SK&A 2016	21.3	21.5	12.9	13.5

Table 5.E.1. ***(continued)***

Practice characteristics at baseline	Data source for characteristic	Track 1		Track 2	
		Mean among CPC+ practices (N = 1,373)	Weighted mean among comparison practices (N = 5,243)	Mean among CPC+ practices (N = 1,515)	Weighted mean among comparison practices (N = 3,783)
3–4 primary care practitioners (%)	SK&A 2016	23.2	24.0	22.4	22.1
5–7 primary care practitioners (%)	SK&A 2016	25.8	25.5	26.0	26.3
8+ primary care practitioners (%)	SK&A 2016	29.8	29.0	38.7	38.1
Practice is multispecialty[b] (%)	SK&A 2016	19.6	20.1	26.2	26.2
Hospital Referral Region price index	CMS's Medicare Geographic Variation data, 2015	1.1	1.1	1.0	1.1
Meaningful EHR use[c] (%)					
Never attested (%)	CMS's Medicare EHR Incentive Program data	8.0	8.5	3.5	3.7
Attested since 2011 or 2012 (%)	CMS's Medicare EHR Incentive Program data	78.9	78.5	88.2	87.9
Attested since 2013 or later (%)	CMS's Medicare EHR Incentive Program data	13.1	13.0	8.3	8.4
Number of Medicare FFS beneficiaries assigned in 2016 per PCP	Mathematica attribution based on SK&A roster	231.0	226.0	197.0	202.0

Source: Mathematica's analysis of baseline practice characteristic data of CPC+ and matched comparison practices.

Note: Because CPC+ is a practice-level model, and to aid computation, we matched using practice-level data rather than beneficiary-level data. However, we analyzed Medicare claims-based outcomes using beneficiary-level data rather than practice-level data, so we show balance statistics to approximate beneficiary-level balance. This approach best reflects the baseline balance in the analytic sample that we used in regression analyses. Specifically, the means in this table represent practice-level means, weighted by the number of Medicare FFS beneficiaries assigned to each practice in 2016.

[a] We define prior transformation experience as CPC Classic or MAPCP participation, or whether the practice is recognized as a medical home by NCQA, TJC, AAAHC, URAC, or a state medical-home recognition program.

[b] Defined as having at least one practitioner, according to SK&A, with a specialty other than general practice, internal medicine, family medicine, or geriatrics.

[c] Defined as having at least one practitioner within the practice who attested to meaningful use under the CMS Medicare EHR Incentive Program.

AAAHC = Accreditation Association for Ambulatory Health Care; ACO = Accountable Care Organization; CAH = critical access hospital; CMS = Centers for Medicare & Medicaid Services; ED = emergency department; EDB = Medicare enrollment database; EHR = electronic health record; FFS = fee-for-service; HCC = hierarchical condition category; MAPCP = Multi-payer Advanced Primary Care Practice Demonstration; MDM = CMS master data management system; NCQA = National Committee for Quality Assurance; PBPM = per beneficiary per month; PCP = primary care practitioner; SSP = Medicare Shared Savings Program; TJC = The Joint Commission; URAC = Utilization Review Accreditation Commission.

5.E.2. Study population and unit of observation in the regression analysis

A. Study population

We used a cross-sectional approach to define the study population, with highly overlapping cross-sections for (1) the baseline year and (2) each year of CPC+. The study population was based on beneficiary attribution (described in Appendix 5.B), and the annual cross-sections of beneficiaries for the baseline year and the intervention period were based on quarterly attribution (see Table 5.E.2 below).

Table 5.E.2. Baseline and intervention year cross-section definitions for study population

	Study population definition
Cross-section	**Beneficiaries attributed to CPC+ or comparison practices at any time during the…**
Baseline	Baseline year (January 1, 2016, to December 31, 2016)
First intervention year	First intervention year (January 1, 2017, to December 31, 2017)
Second intervention year	Second intervention year (January 1, 2018, to December 31, 2018)
Third intervention year	Third intervention year (January 1, 2019, to December 31, 2019)
Fourth intervention year	Fourth intervention year (January 1, 2020, to December 31, 2020)
Fifth intervention year	Fifth intervention year (January 1, 2021, to December 31, 2021)

B. Assignment to the CPC+ or comparison group, based on attribution

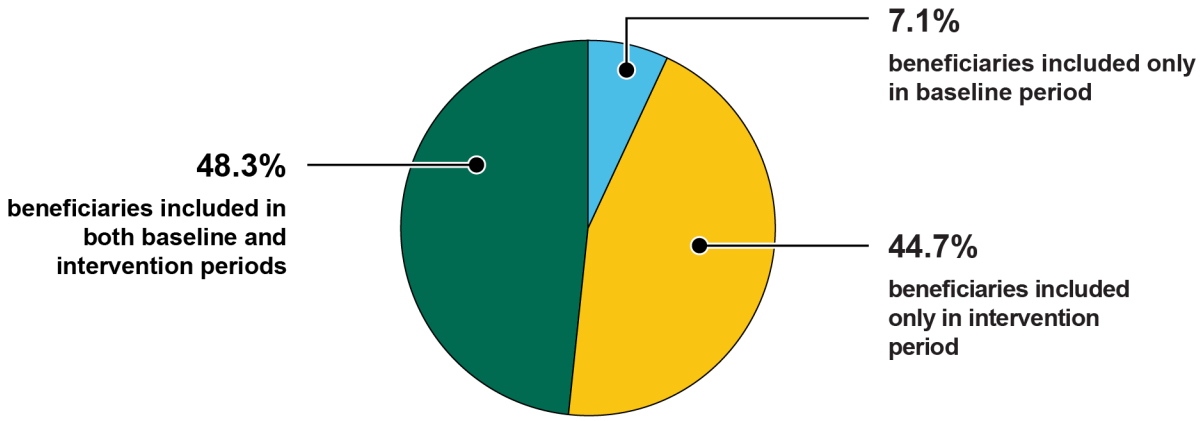
We assigned beneficiaries to the CPC+ or comparison group at two points:

1. For the **baseline period**, we assigned beneficiaries to the CPC+ or comparison group based on the first practice they were attributed to during the baseline period.
2. During the **intervention period**, we assigned beneficiaries to the CPC+ or comparison group based on the first CPC+ or comparison practice they were attributed to during the intervention period; following an intent-to-treat (ITT) approach, we continue to assign the beneficiary to the same practice for the entire intervention period, regardless of whether the beneficiary continued to receive care at that practice as long as they are observable in Medicare Part A and B claims data.

Following these definitions, it is possible for a beneficiary to be in the study population (1) only during the baseline period—for example, if the beneficiary died during the baseline period or was no longer attributed to a CPC+ or comparison practice during the intervention period; or (2) only during the intervention period—for example, if the beneficiary was first attributed to a CPC+ or comparison practice during an intervention year (including people who were new to Medicare). We found that 48.3 percent of beneficiaries were included in both the baseline and intervention periods in our main impact analysis, whereas 7.1 and 44.7 percent, respectively, were included for only the baseline year and only the intervention years (Figure 5.E.1). Because we are retaining beneficiaries in the study population over time (following the ITT approach), as well as adding new beneficiaries to the sample, the sample size during the intervention period will continue to grow as we add more intervention years to the analysis and will include more new beneficiaries compared to the baseline period. Therefore, the percentage of beneficiaries in the full sample—which covers both the baseline and intervention periods—who are only

in the baseline period will fall over time, while the percentage of beneficiaries who are only in the intervention period will increase over time.

Figure 5.E.1. Overlap of beneficiaries in the baseline and intervention periods



Source: Overlap of assigned Medicare FFS beneficiaries in Mathematica's evaluation sample for the five program years and in the year before the start of CPC+ using Medicare claims data from January 2014 to December 2021.

Note: Percentages do not sum to 100 percent due to rounding.

Given the ITT approach to assignment, beneficiaries cannot switch practices *during* the baseline period or *during* the intervention period. This ensures that there is no contamination of the comparison group during the intervention period. However, going from the baseline to the first year of the intervention period, changes in the beneficiary sample at a practice can occur due to:

1. Beneficiaries switching practices—within the CPC+ or comparison group or across groups—since the ITT rule is applied separately in each period. This does not pose a risk of contamination since there was no intervention during the baseline period. Also, practice switches between the baseline and intervention periods are most likely to occur *within* the CPC+ or comparison group, given that we use external comparison regions for matching.
2. Adding beneficiaries who are newly attributed to a CPC+ or comparison practice and found to be eligible.
3. Excluding previously attributed beneficiaries who are no longer eligible (e.g., due to death or enrollment in a Medicare Advantage [MA] plan).

During the intervention period, changes in the beneficiary sample at a practice can occur across years only due to the second and third reasons.

There are two advantages to using an ITT approach for this analysis:

1. It reduces potential biases in impact estimates that could result if CPC+ affects who is attributed to practices over time or which practices are in the sample. For example, through practices' implementation of CPC+ components like care management, enhanced access, and care coordination,

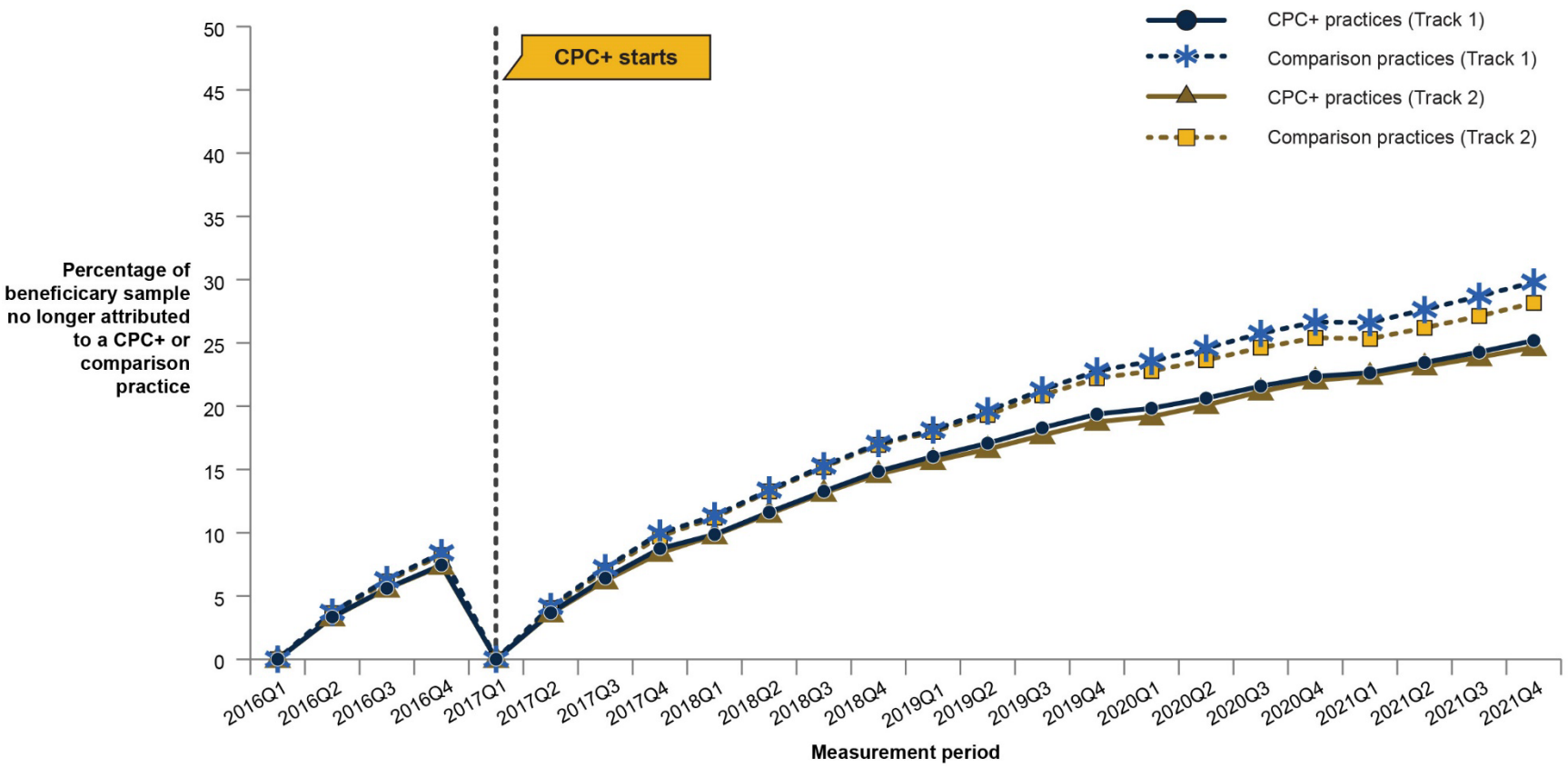
patients in CPC+ practices may be more likely to find a "home" in their CPC+ practice, leading to fewer patients, particularly high-risk patients, switching practices relative to the comparison practices. Thus, in the absence of the ITT approach, we would erroneously estimate that CPC+ increased Medicare expenditures simply because CPC+ practices retained more high-risk patients than the comparison practices. Another example would be if practices stopped treating certain types of beneficiaries due to CPC+ financial incentives, the ITT approach would continue to assign those beneficiaries to the originally attributed CPC+ practices in the following intervention years. CPC+ could also affect whether practices merge, split, or close, for example, by providing enhanced payments.

2. Beneficiaries might continue to benefit from new or improved services they receive from CPC+ practices, even after switching to non-participating practices. A non-ITT approach would miss these effects of CPC+ and potentially attribute them to the non-participating practices.

A disadvantage of the ITT approach is that the estimated impacts of CPC+ could be diluted compared to what would happen if we followed a set of beneficiaries that continuously received care from CPC+ practices. Figure 5.E.2 shows the percentage of beneficiaries who were no longer attributed to a CPC+ or comparison practice during the quarter but were retained after being attributed in a previous quarter, due to the ITT approach. In the first quarter of the baseline period and the first quarter of the intervention period, all beneficiaries in the analytic sample were also originally attributed to a CPC+ or comparison practice by design (since ITT is not applicable in the first quarter of each period). By the last quarter of Program Year (PY) 5 (2021), for both Tracks 1 and 2, about 25 percent of beneficiaries in CPC+ practices were no longer attributed to a CPC+ practice but were still in the research sample; about 30 percent of beneficiaries in Track 1 and 28 percent of beneficiaries in Track 2 were no longer attributed to a comparison practice but were still in the research sample. This finding suggests that, over time, a slightly higher proportion of beneficiaries in CPC+ practices continued receiving billable care from the same type of practices, and therefore continued to be attributed to the same practices, than the proportion of beneficiaries in comparison practices.[64] We conducted a sensitivity analysis for our key outcomes (Medicare expenditures without enhanced payments, acute hospitalizations, and outpatient ED visits) that dropped beneficiaries from the sample when they are no longer attributed to a CPC+ or comparison practice. (See Section 5.E.10 for a more detailed description of this analysis.)

[64] There are many factors that could contribute to this growing difference in the proportion of beneficiaries who remain attributed to CPC+ versus comparison practices, but we cannot fully measure the extent to which the difference is caused by CPC+. CPC+ might make it more likely for beneficiaries to continue to obtain care from the same practice, compared to comparison beneficiaries, due to changes in CPC+ practices including: providing improved patient care due to the care delivery requirements of the model, actively providing and billing for annual wellness visits (which was added as a criterion for getting attributed in 2019; see Appendix 5.B), or continuing to keep their doors open due to the enhanced payments from CPC+. Differences in annual wellness visits appear to explain only a small amount of the CPC+ and comparison differential–for example, the addition of this criterion in 2019 led to a 2.5 percent increase in attributed beneficiaries for CPC+ practices and a 2.3 percent increase for comparison beneficiaries through 2018, leaving a small net differential. Other factors that might contribute to differences between CPC+ and comparison practice beneficiaries being attributed to the same practice could be unrelated to CPC+. For example, there could be selection bias in the model: CPC+ practices presumably would not have applied to CPC+ if they knew they were about to close or their practitioners were about to retire; unfortunately, our evaluation matching design did not include variables such as practitioner age that could have helped mitigate selection bias that leads to differential attrition. Another contributing factor could be data quality issues: since CPC+ practices applied to participate in CPC+, practices in IQVIA rosters that we have identified as the CPC+ practices are less likely to be determined as "erroneous" by IQVIA (as they clean and revise their data) and to disappear from their rosters over time than comparison practices.

Figure 5.E.2. Percentage of beneficiaries in the analytic sample who were no longer attributed to a CPC+ or comparison practice but remained in the research sample due to the ITT approach, by track



Notes: The numbers in this figure represent the percentage of beneficiaries who were no longer attributed to a CPC+ or comparison practice but were retained in the analytic sample due to the ITT sample construction approach. We conduct assignment separately in the baseline and intervention periods. In the first quarter of the baseline period (2016Q1) and in the first quarter of the intervention period (2017Q1), the sample includes only beneficiaries actually attributed during these quarters. In subsequent quarters, beneficiaries remain in the sample even if they are no longer attributed to a CPC+ or comparison practice. Therefore, the percentage of beneficiaries not attributed is zero in 2016Q1 (and then increases over the baseline period) and is zero again in 2017Q1 (and then increases over the intervention period). This figure does not account for attrition among CPC+ practices. That is, beneficiaries attributed to a practice that stopped participating in CPC+ are still considered as being attributed to a CPC+ practice. Approximately 15 percent of CPC+ practices were terminated by CMS, withdrew, or closed during the five years of CPC+.

ITT = intent-to-treat; Q = quarter.

C. Sample size

For Track 1, the main analyses included 1,549,585 unique Medicare FFS beneficiaries served by 1,373 CPC+ practices and 5,347,499 unique beneficiaries served by 5,243 matched comparison practices during either the baseline period or the five program years.[65]

For Track 2, the main analyses included 1,896,880 unique Medicare FFS beneficiaries served by 1,515 CPC+ practices and 4,507,499 unique beneficiaries served by 3,783 matched comparison practices during either the baseline period or the five program years.

D. Unit of observation

The unit of observation in the regressions for all claims-based outcomes (other than the 30-day readmissions, unplanned acute care outcomes, and comprehensiveness of care outcomes) is the beneficiary-year. Each beneficiary has observations for as many years as the beneficiary remains in the sample (as defined above) and can still be observed in claims. Specifically, to be observed, a beneficiary assigned to a practice for the baseline or the intervention period had to be alive, have both Medicare Part A and B FFS coverage with Medicare as the primary payer, and not be covered under a Medicare Advantage or other Medicare health plan.[66] Medicare beneficiaries who were dually eligible for Medicaid will be attributed as long as they meet the other eligibility requirements.

For the 30-day readmissions and the unplanned acute care after hospitalization outcomes, for which we only included beneficiaries who had at least one eligible hospital discharge in a year, the unit of analysis is the index hospital discharge, rather than the beneficiary. So, for example, a beneficiary who has two index hospital discharges in a year has two observations in that year, one for each discharge.[67] Similarly, for the unplanned acute care after an emergency department (ED) visit or an observation stay outcome, the unit of analysis is the index ED visit or observation stay.

If CPC+ practices are more effective in keeping beneficiaries out of the hospital or the emergency room, the relative severity of index discharges (including index hospital discharges and index ED visits or observation stays) could rise for the CPC+ group compared to the comparison group over time and might include discharges that are more likely to result in a readmission or an unplanned acute care event. This change in the relative severity of index discharges could lead to higher readmission or unplanned acute care rates in the CPC+ group. To address this issue, we conducted a sensitivity test using readmission and unplanned acute care measures calculated at the beneficiary level. For this test, we include all beneficiaries in the sample—even those without any index hospitalizations, or index ED visits or observation stays.

---

[65] After accounting for weights that adjust for matching and time observed in Medicare FFS, the effective sample sizes in the main analysis for the baseline period are about 95 percent of the actual sample size for the CPC+ sample in both Tracks 1 and 2 and about 45 and 40 percent of the actual sample size for comparison sample in Tracks 1 and 2, respectively. The ratio of the effective sample size to the actual sample size is higher for the CPC+ sample than for the comparison sample because the matching weight, by default, was equal to 1 for the CPC+ sample.

[66] As we describe in Appendix 5.B, we apply an additional criterion for a beneficiary not being incarcerated when we identify attributed patients, following CMS's approach to patient attribution. Once we attribute a patient to a CPC+ or comparison practice based on all criteria in the attribution algorithm, the final analysis ignores the "not incarcerated" requirement in identifying the number of FFS eligible months for patients.

[67] A readmission could qualify as an index stay if it meets the eligibility criteria for an index hospital admission.

For the comprehensiveness of care outcomes, for which the study population is primary care practitioners (as defined by the National Provider Identifier [NPI]) who were affiliated with CPC+ or comparison practices at baseline or at the start of the first program year (2017),[68] the unit of observation in the regressions is the practitioner-year. If a practitioner was affiliated with multiple practices (within our sample of CPC+ and comparison practices) in a year, we randomly assigned that practitioner to a single practice. Approximately 3.3 to 3.6 percent of the practitioners were affiliated with multiple practices for any given baseline or program year.

5.E.3. Model specification

In this section, we describe both the difference-in-differences model used for most outcomes and the straight-difference model (defined below) used for the telehealth and mortality outcomes. We note key differences in the estimation of the difference-in-differences model for the 30-day readmissions, unplanned acute care outcomes, and comprehensiveness of care outcomes in Sections 5.E.4, 5.E.6, and 5.E.7.

A. Difference-in-differences model

We estimated the impact of CPC+ by using difference-in-differences regressions. Specifically, for all our beneficiary-level outcomes except for telehealth and mortality, we compared the difference in mean outcomes between beneficiaries assigned to CPC+ and comparison practices during (1) the baseline year before CPC+ (2016) and (2) each intervention year of CPC+ (Years 1 through 5), while controlling for beneficiary characteristics at baseline, COVID-19-related controls, and practice-level fixed effects. Since the impact analysis includes PYs 4 and 5 or calendar years 2020 and 2021, it was important to account for any differences in how the COVID-19 pandemic unfolded in CPC+ versus comparison regions. Therefore, we included COVID-19-related controls in the impact analysis, based on the detailed claims-based COVID checks that are described in Appendix 5.D. The beneficiary-level controls, COVID-19-related controls, and the practice fixed effects help to (1) adjust for beneficiary risk; (2) mitigate potential bias in PYs 4 and 5 CPC+ impact estimates due to differences between CPC+ and comparison regions in the timing, severity, and effects of COVID-19 on mortality and health care use; (3) improve the precision of the model; and (4) account for any remaining imbalance in beneficiary and practice characteristics, including unmeasured and time-invariant practice characteristics at baseline.

In Equation (5.E.1), let *i* index the beneficiary, *j* index the practice, and *t* index time, where *t* ranges from 0 to 5, with 0 denoting the baseline year. Given the study population and unit of observation defined above, for the main regression analyses we estimated difference-in-differences regression models of the following form, with one regression for each outcome:

$$(5.E.1)\quad y_{ijt} = \alpha + \beta X_{it} + \gamma_t p_t + \theta_t z_j p_t + \delta_t C_{jt} p_{t\{t=2020,2021\}} + b_j + \varepsilon_{ijt},$$

where

[68] We applied a physician-oriented ITT approach during the intervention period, that is, we kept following a physician once they were identified to be at a CPC+ or comparison practice in the first program year (2017). We did not use a cross-sectional approach to identify the physician sample during the intervention period, that is, we did not add practitioners who joined a practice or remove practitioners who left a practice after the first program year, because there was an influx of primary care clinicians into practice sites due to the change in practice roster data source in 2019 from SK&A to OneKey.

y_{ijt} represents a claims-based outcome variable for beneficiary *i*, in practice *j*, in year *t*. Outcome variables include Medicare expenditures and measures of utilization such as hospitalizations. Table 5.C.1 in Appendix 5.C lists the outcomes.

X_{it} is a vector of characteristics of beneficiary *i* measured at the start of the baseline period for baseline observations, and at the start of the intervention period for intervention period observations. For example, beneficiary characteristics include demographics (age, race, and gender), variables capturing Medicare and Medicaid eligibility (that is, original reason for Medicare eligibility, and dual Medicare-Medicaid status), and hierarchical condition category (HCC) score. We also include beneficiary characteristics like HCC score interacted with the year indicators (from PY 2 onward) to account for possible changes in the relationship between the characteristic measured at the start of CPC+ and outcomes. We describe covariates in more detail in Section 5.E.6 below.

p_t (for "post") is an intervention-period indicator that takes the value of 1 during any intervention year, and 0 otherwise.

z_j is a binary indicator of intervention status or of being in a CPC+ practice; the indicator takes the value of 1 if practice *j* is a CPC+ practice, and is otherwise 0. The main effect of this indicator is not identified in this equation since it is collinear with the practice fixed effects.

C_{jt} is a vector of COVID-19-related controls including excess deaths in the state-hospital referral region (HRR), Pandemic Vulnerability Index in the county, Government Response Index in the state, and Social Vulnerability Index in the census tract of each practice *j* in year *t*. We include COVID-19-related controls interacted with the contemporaneous year indicator (2020 or 2021) to account for potential effects of COVID-19 on outcomes in calendar years 2020 and 2021, respectively.[69]

b_j is a practice-level fixed effect for practice *j*, which controls for all time-invariant practice characteristics.

ε_{ijt} is the idiosyncratic error term. It represents unexplained variability in the outcome variable for beneficiary *i*, in practice *j*, during period *t*.

B. Straight-difference model

For telehealth service use and expenditures, as well as mortality, we estimated the impact of CPC+ by using straight-difference regressions, comparing the difference in mean outcomes between beneficiaries assigned to CPC+ and comparison practices during a specific observation period. We used the straight-difference model instead of the difference-in-differences model for telehealth outcomes since the use of these services was close to zero at baseline. In other words, the mean outcome in any intervention year for the CPC+ or comparison group is similar to the change in the mean outcome from baseline to that intervention year for telehealth services. In addition, we only modeled the telehealth outcomes in PYs 4

[69] The main effects of these COVID-19-related regional controls are not identified in this equation because the model includes practice fixed effects.

and 5 because the use of these services was also close to zero in the first three intervention years. Since the probability of dying increases with the length of the observation period, we decided to model mortality over fixed lengths of follow-up (for example, 12, 24, 36, 48, and 60 months) during the intervention period with a straight-difference model. We controlled for beneficiary and practice characteristics at baseline, and COVID-19-related controls (for regressions that include observations in PYs 4 and 5).

In Equation (5.E.2), let *i* index the beneficiary and *j* index the practice. For the telehealth and mortality outcomes, we estimated straight-difference regression models of the following form, with a separate regression for each outcome in each intervention year[70]:

$$(5.\mathrm{E}.2)\quad y_{ij} = \alpha + \beta X_i + \pi z_j + \rho C_j + \mu D_j + \varepsilon_{ij},$$

where

y_{ij} represents a telehealth or mortality outcome variable for beneficiary *i* in practice *j*. Telehealth outcome variables include the proportion of ambulatory primary care visits and expenditures that are non-face-to-face with primary care practitioners in PY 4 or PY 5. Mortality outcome variables include whether a patient died within 12 months since the start of the baseline period, and whether a patient died within 12, 24, 36, 48, or 60 months since the start of the intervention period (that is, by the end of PYs 1 through 5). We also looked at the fraction of days alive across 12 months since the start of the baseline period and fraction of days alive across the 12, 24, 36, 48, and 60 months since the start of the intervention period. Table 5.C.1 in Appendix 5.C lists the telehealth outcomes and Section 5.C.1.G in Appendix 5.C lists the mortality outcomes.

X_i is a vector of baseline characteristics of beneficiary *i* as in Equation (5.E.1). For modeling telehealth outcomes, we also control for baseline Medicare expenditures and use of selected services to account for differences in health care utilization between beneficiaries assigned to CPC+ versus comparison practices before the start of CPC+. We describe the baseline Medicare expenditures and service use control variables in more detail in Section 5.E.6 below.

z_j is a binary indicator of being in a CPC+ practice as in Equation (5.E.1); the indicator takes the value of 1 if practice *j* is a CPC+ practice, and is otherwise 0.

C_j is a vector of COVID-19-related controls as in Equation (5.E.1). We included contemporaneous COVID-19-related controls for examining telehealth outcomes in PY 4 or PY 5. We also included COVID-19-related controls in 2020 and 2021for examining the 48-month and 60-month mortality outcomes that include data through PYs 4 and 5, respectively.

[70] Since the use of telehealth varied widely between 2020 and 2021 (likely due to changes in intensity of the pandemic, regulatory policies, and practice adaptability), we estimated separate regressions in each intervention year (2020 and 2021) for telehealth outcomes to allow for the correlations between beneficiary/practice characteristics and outcomes to vary by year.

D_j is a vector of baseline characteristics of practice *j*. We describe practice-level control variables in more detail in Section 5.E.6 below.

ε_{ij} is the idiosyncratic error term. It represents unexplained variability in the telehealth or mortality outcomes for beneficiary *i* in practice *j*.

5.E.4. Model output and interpretation of key coefficients

In Equation (5.E.1) (difference-in-differences model), the intervention period-specific coefficients (γ_t) capture changes experienced by the comparison group in each intervention-period interval. Note that, instead of assuming a linear time trend, we allowed the coefficients to vary for each interval. The set of interaction terms ($\theta_t z_j p_t$) captures the difference in outcomes between the CPC+ and comparison groups for each intervention-period interval relative to that difference in the baseline period, adjusting for differences in (observed) beneficiary and (observed and unobserved) practice characteristics that remain after matching. Thus, the θ_t coefficients are the interval-specific impact estimates that capture whether CPC+ made a difference to an outcome of interest.

By estimating Equation (5.E.1) for the impact analysis in this report, we obtained an estimate of θ_t for each year of CPC+, as well as regression-adjusted means for baseline and intervention years, by intervention status. In addition to the model specified by Equation (5.E.1), we estimated an alternative model that controls for the same set of variables as in Equation (5.E.1) but assumed a constant impact θ across the entire intervention period, providing an average impact estimate across the five intervention years.

Table 5.E.3 illustrates how the parameter estimates from Equation (5.E.1) can be used to obtain the regression-adjusted CPC+ and comparison group means for the baseline year and each intervention year, along with the difference-in-differences impact estimates for Years 1 through 5. Because we use practice fixed effects, the main effect of intervention status, or the coefficient on the indicator for being in a CPC+ practice (the parameter φ in Table 5.E.3) cannot be estimated by Equation (5.E.1). Therefore, in our report, we use the following approach to show CPC+ and comparison group means in tables reporting difference-in-differences estimates. We show the actual, unadjusted CPC+ means at baseline and each intervention year. For the comparison group, we show the actual, unadjusted mean at baseline and the adjusted mean in each intervention year. We obtained this adjusted mean by subtracting the regression-adjusted difference between the CPC+ and matched comparison groups in each year (obtained from the difference-in-differences model) from the unadjusted CPC+ mean in that same year. We also calculated percentage impacts relative to what the CPC+ mean would have been in an intervention year in the absence of CPC+—that is, the unadjusted CPC+ mean minus the impact estimate.

The general model specification, output, and interpretation of key coefficients for the 30-day readmissions, unplanned acute care outcomes, and comprehensiveness of care outcomes are the same as for the beneficiary-year level outcomes, except that the model is specified at the discharge level for the former and the practitioner-year level for the latter.

In Equation (5.E.2) (straight-difference model), the coefficient π on the CPC+ practice indicator is the impact estimate that captures whether CPC+ made a difference to a telehealth outcome or to a period-specific mortality outcome.

Table 5.E.3. Impact estimates and CPC+ and comparison group means based on a linear regression from Equation (5.E.1): a stylized representation

Year	CPC+ group mean	Comparison group mean	Difference between CPC+ and comparison means	Difference-in-differences impact estimate
Baseline year ($t=0$) [reference period]	$\alpha+(\varphi)$	α	(φ)	N/A
First intervention year ($t=1$)	$\alpha+(\varphi)+\gamma_1+\theta_1$	$\alpha+\gamma_1$	$(\varphi)+\theta_1$	θ_1
Second intervention year ($t=2$)	$\alpha+(\varphi)+\gamma_2+\theta_2$	$\alpha+\gamma_2$	$(\varphi)+\theta_2$	θ_2
Third intervention year ($t=3$)	$\alpha+(\varphi)+\gamma_3+\theta_3$	$\alpha+\gamma_3$	$(\varphi)+\theta_3$	θ_3
Fourth intervention year ($t=4$)	$\alpha+(\varphi)+\gamma_4+\theta_4$	$\alpha+\gamma_4$	$(\varphi)+\theta_4$	θ_4
Fifth intervention year ($t=5$)	$\alpha+(\varphi)+\gamma_5+\theta_5$	$\alpha+\gamma_5$	$(\varphi)+\theta_5$	θ_5

Notes: To highlight the key coefficients in Equation (5.E.1), we exclude the coefficients on beneficiary characteristics, practice characteristics, and COVID-19-related controls in the expressions for the CPC+ and comparison group means in this table. The parameter φ in the table denotes the main effect of intervention status, or a coefficient on the indicator for being in a CPC+ practice. This term is not included in Equation (5.E.1); it cannot be directly estimated because the model includes practice fixed effects. We include this term in this table to illustrate the difference-in-differences approach, but we show it in parentheses since we do not obtain an estimate of it. This parameter is differenced out in obtaining the impact estimate.

5.E.5. Model estimation

A. Separate regressions by track and by Medicare Shared Savings Program (SSP) status

For each Medicare claims-based outcome of interest, we estimated six separate regressions for our main analysis. We estimated impacts separately for Track 1 and Track 2, given that participating practices face track-specific requirements, payments, and incentives, which may yield very different impacts. Within each track, in addition to an overall estimate of CPC+, we also estimated impacts separately by SSP

participation status at the start of CPC+ (January 1, 2017).[71,72] For selected outcomes, we also estimated impacts separately for other key subgroups, by including additional interaction terms in the regression, as we describe below in Section 5.E.9.

B. Linear regression

For Medicare expenditures, and for any other continuous outcomes (which include service use outcomes, continuity of care outcomes, comprehensiveness of care, length of hospice, and the composite low-value services outcome), we estimated Equations (5.E.1) and (5.E.2) as a linear regression. We also used linear regressions for all binary outcomes (which include unplanned readmissions and unplanned acute care, any hospice use, mortality, receipt of recommended services for beneficiaries with diabetes and for breast cancer screening, appropriate use of medications, long-term opioid use, and potential opioid overuse). An alternative approach would have been to use generalized linear models to account for the distinctive distributional features of service use outcomes and use logistic regression for binary outcomes. However, from the perspective of computational feasibility, nonlinear models were expected to be much more resource- and time-intensive given the large sample sizes. Also, we were more likely to experience problems with model convergence with a nonlinear model, especially when using a specification with practice fixed effects, due to features in the data (for example, a binary outcome being equal to zero or one for all beneficiaries in a practice or for all beneficiaries with a certain combination of characteristics). Therefore, our preferred approach was to estimate linear regressions for all outcomes. We tested how much the choice of functional form might influence the results of our impact evaluation, and we found we obtained nearly identical point estimates of the difference-in-differences impacts using either linear or nonlinear models.[73]

C. Non-independence

All regressions accounted for non-independence across observations within the same practice using standard error estimates clustered at the practice level. Although this approach yields consistent standard error estimates, we considered alternatives for two reasons. First, because there is much stronger correlation across repeated observations from the same beneficiary than among beneficiaries receiving care from the same practice, we tested whether explicitly accounting for beneficiary-level clustering would improve standard error estimates. Second, we tested whether including fixed or random effects at the beneficiary or practice level could help guard against omitted-variable bias by controlling for any

[71] Practices may change their SSP status over the course of CPC+, but we do not control for this change, because participation in CPC+ may cause a practice to participate in (or drop out of) SSP.

[72] An alternative to estimating separate models by SSP participation status is to use a triple differences estimation approach, where the coefficient on the triple interaction term for SSP participation, participation in CPC+, and the intervention period dummy would provide the impact estimate for SSP practices. Ideally, we would also allow the effect of beneficiary demographics and other practice characteristics (fixed effects) to vary by SSP participation status. However, allowing for the effect of each of the model covariates to vary by SSP participation status could make the estimation unwieldy. Therefore, we estimated impacts using separate regressions for SSP practices and non-SSP practices within each track.

[73] In a sensitivity analysis comparing inference from two models that were identical except that one was a linear regression and the other was a zero-inflated negative binomial model, we found that across the four years of CPC Classic, the two approaches gave nearly identical point estimates of the difference-in-differences impact for a count variable of number of hospitalizations. The linear model's standard errors around those point estimates were about 10 percent larger than those from the zero-inflated negative binomial model. Therefore, using a linear model should provide us with point estimates similar to those from a more complex, maximum likelihood model, but slightly more conservative standard errors, potentially lowering the likelihood that a small to moderate-size effect is considered statistically significant.

time-stable unmeasured beneficiary- or practice-level confounders. The detailed testing methods and results are in Appendix 3.O of the evaluation design report (Orzol et al. 2022). We found that a model with practice-level fixed effects and standard error estimates clustered at the practice level provided the best performance in terms of the mean squared error of the difference-in-differences point estimate and the coverage of the confidence interval around this estimate.[74] Therefore, we adopted this approach to account for non-independence.

D. Interpretation

We calculated all impact estimates at the beneficiary-year level (or the discharge-year level for readmissions and unplanned acute care outcomes or the practitioner-year level for comprehensiveness of care outcomes), but we sometimes describe them as differential changes experienced by CPC+ versus comparison practices in our discussion of results, because CPC+ is a practice-level model.

We used regression output to calculate p-values for statistical inference. We used two-tailed tests with $p < 0.10$ as the threshold of statistical significance. Although we did not apply any formal multiple comparison corrections (many of which are known to be overly conservative), our approach to interpreting impact estimates aimed to avoid "false positives" (Peterson et al. 2018). To minimize the probability of mistaking noise for signal when examining impacts, we combined evidence from p-values with evidence from subgroup analyses, related outcomes, sensitivity tests, and the implementation analysis to reinforce or discount the interpretation of observed results.

5.E.6. Control variables

A. Control variables for most outcomes

The regressions for most outcomes (other than discharge-level outcomes, comprehensiveness of care, telehealth, and mortality) controlled for beneficiary characteristics, COVID-19-related controls, and practice fixed effects. The beneficiary-level control variables included demographics (age categories, race and ethnicity, and gender), original reason for Medicare entitlement, dual eligibility status, and HCC score (Table 5.E.4). For comprehensive risk adjustment, the regression additionally includes indicators for specific chronic conditions that are prevalent in the CPC+ sample, defined by applying the HCC or Chronic Conditions Warehouse (CCW) algorithm on Medicare claims (see Appendix 5.C for more information on how we selected the HCCs to include as controls in most regressions; also see Appendix 5.G for additional HCCs used as control variables in the regressions for the long-term opioid use and potential opioid overuse outcomes). We also include an indicator that the HCC score was calculated using only demographic information as a control variable.[75] We included interactions of HCC score and chronic conditions with indicators for the second and each subsequent intervention year to account for possible

[74] Although practice fixed effects account for part of the within-practice correlation in outcomes, they do not account for such correlation completely. Specifically, practice fixed effects assume a fixed degree of correlation between any two observations from the same practice. In reality, however, there could be differences in the degree of correlation arising due to different beneficiaries being in the same practice versus correlation in outcomes over time for the same beneficiary in that practice (autocorrelation). Also, practice fixed effects do not account for heteroscedasticity. Therefore, using standard error estimates clustered at the practice level on top of practice fixed effects is likely to provide a more accurate estimate of the standard error for the impact estimates.

[75] HCC scores are calculated on the basis of demographic characteristics only when claims data are not observed for a beneficiary and may not reflect the beneficiary's actual risk. This generally happens when the beneficiary is new to Medicare FFS.

changes in the relationship between HCC scores and chronic conditions (measured at the start of CPC+) and outcomes (measured after the first intervention year). For observations in the baseline period, beneficiary-level control variables were measured directly before the start of the yearlong baseline period (based on data from calendar year 2015). For observations in the intervention period, beneficiary-level control variables were measured directly before the start of CPC+ (based on data from calendar year 2016). We did not update the beneficiary characteristics over the intervention period because CPC+ could affect the observed beneficiary characteristics.

Given that we used a difference-in-differences approach, we did not include as control variables Medicare service use or expenditures during the baseline period, as is often done in a cross-sectional analysis. These baseline outcomes are the dependent variable for the baseline observations in our model and, therefore, cannot be viewed as independent of the error term.

COVID-19-related controls were included to mitigate potential bias due to regional differences in the timing, severity, and effects of COVID-19, and behavioral responses to COVID-19 during the fourth and fifth intervention years. COVID-19-related controls include excess deaths in the state-HRR, Pandemic Vulnerability Index in the county, Government Response Index in the state, and Social Vulnerability Index in the census tract of each practice (Table 5.E.4). We interacted each year-specific COVID-19-related variable with the contemporaneous year indicator (2020 or 2021). (See Appendix 5.D for more information on how the COVID-19-related control variables were created.)

For the composite low-value services outcome, which consists of 31 individual low-value services, we additionally included binary indicators for 16 individual service qualifications as control variables.[76] These indicators will account for potential differences between CPC+ and comparison groups in the proportion of beneficiaries qualifying for any measure. Given their similarity to HCCs, to avoid collinearity, we excluded the chronic condition controls for specific HCCs from the low-value services regressions, while retaining the controls for HCC score. For the long-term opioid use and potential opioid overuse outcomes, we additionally controlled for changes in state opioid policies, in order to account for potential confounding due to differential changes in state-level opioid policies and practices over time between CPC+ and comparison groups. (See Appendix 5.G for more information on the state-level opioid policy variables used as covariates.)

The practice fixed effects are indicators or dummy variables—one for each practice in the CPC+ and comparison groups. Including these effects controls for any inherent, time-invariant differences between the CPC+ and comparison practices—whether such differences are observed or unobserved. Including practice fixed effects ensured that we accounted for any remaining imbalance in the practice-level variables used in matching, and in any other unmeasured practice characteristics at baseline, when obtaining the difference-in-differences impact estimates. We did not incorporate changes over time in observed practice characteristics as control variables, because CPC+ could affect practice characteristics.

[76] The 16 binary indicators for individual service qualifications included patients with chronic kidney disease receiving dialysis, patients with chronic kidney disease not receiving dialysis, patients aged 75 years or older, women aged 65 years or older, men aged 75 years or older, patients with osteoporosis, patients with deep vein thrombosis, patients with hypothyroidism, patients with a diagnosis of rheumatoid arthritis, patients undergoing selected surgeries for preoperative pulmonary function testing, patients undergoing selected surgeries for preoperative echocardiography and stress testing, patients with syncope diagnosis, patients with fasciitis diagnosis, patients with ischemic heart disease, patients with hypertension, and patients who were hospitalized with a non-surgical Medicare Severity Diagnosis Related Group.

B. Control variables for discharge-level outcomes

As we noted previously, our analyses for readmissions and unplanned acute care outcomes are at the discharge-year (rather than beneficiary-year) level. Therefore, the difference-in-differences regressions for these outcomes included some additional control variables. Specifically, we included indicators for conditions identified in inpatient or ED episodes of care during the 12 months before the index admission or the index ED visit or observation stay as well as those present at the index event (there are 31 such condition categories for this analysis). Given their similarity to HCCs, to avoid collinearity, we excluded the chronic condition controls for specific HCCs from the readmission and unplanned acute care regressions, while retaining the controls for HCC score. We also controlled for whether the principal diagnosis or procedure associated with the index discharge is best classified as (1) medicine, (2) surgery/gynecology, (3) cardiorespiratory, (4) cardiovascular, or (5) neurology.[77]

C. Control variables for telehealth and mortality outcomes

As we noted previously, our analyses for telehealth and mortality outcomes use a straight-difference model instead of the difference-in-differences model. We still adjusted for the beneficiary-level control variables in each regression and COVID-19-related controls (for regressions that include observations in PYs 4 and 5) as in the difference-in-differences models. However, we did not include any beneficiary characteristics interacted with the year indicators, which cannot be estimated because there is only one year included in each model. Also, the regressions for telehealth and mortality outcomes controlled for baseline practice-level control variables (Table 5.E.4) instead of practice fixed effects (the treatment effect cannot be identified in the straight-difference model that includes practice fixed effects due to collinearity). For the telehealth outcomes, to adjust for differences in health care utilization among beneficiaries attributed to the CPC+ and comparison practices at baseline, we additionally included the average monthly Medicare expenditures, annualized number of acute hospitalizations, outpatient ED visits, and ambulatory primary care visits, and an indicator for whether baseline Medicare expenditures and services utilization were missing.

D. Control variables for comprehensiveness of care outcomes

As we noted previously, our analysis for comprehensiveness-of-care outcomes is at the practitioner-year (rather than beneficiary-year) level. Therefore, the regression for these outcomes includes control variables at the practitioner level instead of at the beneficiary level. Specifically, we controlled for a practitioner's age (<31, 31–50, and >50 years), gender (binary indicator for male), and primary specialty (family, general, internal, pediatric, and geriatric practice).[78] We also included practice fixed effects and COVID-19-related controls, as in the difference-in-differences models.

[77] The 31 condition categories for the Medicare analysis include a range of diagnoses or risk factors, such as severe infection, metastatic cancer/acute leukemia, diabetes mellitus, end-stage liver disease, drug and alcohol disorders, congestive heart failure, chronic obstructive pulmonary disease, ulcers, cardiorespiratory failure or cardiorespiratory shock, acute renal failure, transplants, hip fracture/dislocation, and more. Our approach was based on reviewing standard models in the literature for risk-adjusting the likelihood of readmission, although it differed from other models in that we did not estimate a separate readmission or unplanned acute care equation for each of the specialty cohorts (medicine, surgery, cardiorespiratory or cardiovascular, or neurology), given our goal of estimating the impact of CPC+ on the risk of all unplanned readmissions or acute care use. The lookback period for these conditions is one to three years, depending on the condition, as specified in the Yale algorithm (YNHHSC/CORE 2019).

[78] Practitioner-level control variables were measured either directly before the start of CPC+ (for the intervention-period observations) or directly before the start of the baseline period (for the baseline-period observations).

Table 5.E.4. Control variables used in the impact analyses for most outcomes

Characteristic	Variables
Beneficiary-level control variables[a]	
Demographics	Age categories < 65 65–74 (reference category) 75–84 ≥ 85 Race/ethnicity[b] Non-Hispanic White (reference category) Non-Hispanic Black Hispanic All other/unknown Gender (binary indicator for male)
Original reason for Medicare eligibility	Original Medicare eligibility categories Age (reference category) Disability only ESRD only or ESRD with disability
Dual eligibility	Indicator for dual status (where dual is defined as those with full or partial Medicaid benefits according to Master Beneficiary Summary File)
Chronic conditions	**HCCs**[c] HCC 8 – Metastatic Cancer and Acute Leukemia HCC 18 – Diabetes with Chronic Complications HCC 21 – Protein-Calorie Malnutrition HCC 22 – Morbid Obesity HCC 23 – Other Significant Endocrine and Metabolic Disorders HCC 85 – Congestive Heart Failure HCC 96 – Specified Heart Arrhythmias HCC 106 – Atherosclerosis of the Extremities with Ulceration or Gangrene HCC 111 – Chronic Obstructive Pulmonary Disease HCC 173 – Traumatic Amputations and Complications HCC 186 – Major Organ Transplant or Replacement Status HCC 40 or 47 – Rheumatoid Arthritis and Inflammatory Connective Tissue Disease or Disorders of Immunity HCC 46 or 48 – Severe Hematological Disorders, or Coagulation Defects and Other Specified Hematological Disorders HCC 54 or 55 – Drug/Alcohol Psychosis or Dependence HCC 57 or 58 – Schizophrenia or Major Depressive, Bipolar, and Paranoid Disorders HCC 70 or 71 – Quadriplegia or Paraplegia HCC 80 or 82 – Coma, Brain Compression/Anoxic Damage or Respirator Dependence/Tracheostomy Status HCC 86, 87, or 88 – Acute Myocardial Infarction, Unstable Angina and Other Acute Ischemic Heart Disease, or Angina Pectoris HCC 99 or 100 – Cerebral Hemorrhage, or Ischemic or Unspecified Stroke HCC 107 or 108 – Vascular Disease, with Complications HCC 157 or 158 – Pressure Ulcer of Skin with Necrosis Through to Muscle, Tendon, or Bone; or of Skin with Full Thickness Skin Loss **Chronic Conditions Warehouse (CCW) indicator** Alzheimer's disease or dementia HCCs and CCW indicator interacted with follow-up year from second follow-up year onward

Table 5.E.4 ***(continued)***

Characteristic	Variables
Risk score	**HCC score** Indicator for whether HCC score was assigned a new enrollee HCC score, i.e., HCC score was calculated on the basis of demographic characteristics only HCC score interacted with follow-up year from second follow-up year onward Indicator for being assigned a new enrollee HCC score interacted with follow-up year from second follow-up year onward
COVID-19-related controls[d]	
Excess deaths	Monthly excess deaths in the state-HRR averaged during each wave[e] of the pandemic, interacted with the contemporaneous year indicator (2020 or 2021) Maximum monthly excess deaths in the state-HRR in 2020 and 2021 interacted with the contemporaneous year indicator (2020 or 2021) Indicator for the wave[e] that the maximum value occurred (reference: wave 1) interacted with the contemporaneous year indicator (2020 or 2021)
Pandemic Vulnerability Index[f]	Monthly Pandemic Vulnerability Index for each county averaged during each wave[f] of the pandemic interacted with the contemporaneous year indicator (2020 or 2021)
Government Response Index[g]	Government Response Index in the state averaged across 2020 and 2021 interacted with the contemporaneous year indicator (2020 or 2021)
Social Vulnerability Index[h]	Social Vulnerability Index in the census tract in 2018 interacted with year indicators for 2020 and 2021, separately[i]
Practice-level control variables[j]	
Practice characteristics	Number of primary care practitioners: 1–2 primary care practitioners (reference category) 3–5 primary care practitioners 6+ primary care practitioners Indicator for whether practice is multispecialty[k] Indicator for hospital ownership or health system management or ownership **Indicator for any nursing practitioner or physician assistant in the practice** Meaningful EHR use[l] Never attested Attested since 2011 or 2012 (reference category) Attested since 2013 or later Indicator for participation in prior primary care transformation activities[m] Indicator for participation in SSP ACO as of January 1, 2017 SSP track Medicare Advantage penetration in the practice's county Median household income in the practice's county Percentage of persons in poverty in the practice's county Percentage with college degree in the practice's county Indicator for health professionals (primary care) shortage area in the practice's county **Hospital beds per 10,000 population in the practice's county** Quartile 1 (reference category) **Quartile 2** **Quartile 3** **Quartile 4**

Table 5.E.4 *(continued)*

Characteristic	Variables
Practice characteristics *(continued)*	Urbanicity of practice's county Rural Suburban Urban (reference category) HRR price index **Census statistical region** Northeast Midwest (reference category) South West

[a] Beneficiary-level control variables were measured either directly before the start of CPC+ (for the intervention-period observations) or directly before the start of the yearlong baseline period (for the baseline-period observations). The yearlong baseline period is 2016 for the practices that started CPC+ in 2017.

[b] We controlled for race/ethnicity with imputed race and ethnicity data, using a methodology called Medicare Bayesian Improved Surname Geocoding (MBISG 2.1). The set of MBISG race/ethnicity variables included imputed probabilities that each beneficiary is White, Black, Hispanic, or other. These probabilities, which incorporated administrative data, surname, and residential location, are strongly predictive of self-reported race and ethnicity (Haas et al. 2019).

[c] We selected a small subset—21 of the 87 HCCs created by the HCC model—for inclusion as control variables. Of the 87 total HCCs, 79 came from the version 22 2017 HCC model and 8 came from the version 21 2017 ESRD model. We selected the 21 HCCs in the subset based on the relative weight of specific HCCs in the HCC score calculation, as well as their prevalence in our analysis sample. We also included an indicator for Alzheimer's disease or dementia from the Chronic Conditions Warehouse (to ensure consistency with CMS's approach for identifying high-risk, Tier 5 beneficiaries in Track 2 of CPC+).

[d] See Appendix 5.G for more information on how the COVID-19-related control variables are created.

[e] We defined a total of seven waves of the COVID-19 pandemic based on trends in excess deaths: three waves in 2020 and four waves in 2021. Specifically, the three waves in 2020 included March–May (wave 1), June–September (wave 2), and October–December (wave 3). The four waves in 2021 included January–February (wave 4; continued from 2020), March–May (wave 5). June–October (wave 6), and November–December (wave 7).

[f] Data source: National Institute of Environmental Health Sciences, North Carolina State University, and Texas A&M University.

[g] Data source: The Oxford Covid-19 Government Response Tracker.

[h] Data source: Centers for Disease Control and Prevention.

[i] We used the 2018 Social Vulnerability Index, the latest year for which the index is available, rather than the 2016 (baseline) version of the index to capture social vulnerability as close to the pandemic period as possible.

[j] Practice-level control variables were only included in regressions for the telehealth and mortality outcomes.

[k] Defined as having at least one practitioner, according to SK&A, with a specialty other than general practice, internal medicine, family medicine, or geriatrics.

[l] Defined as having at least one practitioner within the practice who attested to meaningful use under the CMS Medicare EHR Incentive Program.

[m] We define prior transformation experience as CPC Classic or MAPCP participation, or whether the practice is recognized as a medical home by NCQA, TJC, AAAHC, URAC, or a state medical-home recognition program.

AAAHC = Accreditation Association for Ambulatory Health Care; ACO = Accountable Care Organization; CMS = Centers for Medicare & Medicaid Services; EDB = Medicare enrollment database; EHR = electronic health record; ESRD = end-stage renal disease; HCC = hierarchical condition category; HRR = hospital referral region; MAPCP = Multi-payer Advanced Primary Care Practice Demonstration; NCQA = National Committee for Quality Assurance; SSP = Medicare Shared Savings Program; TJC = The Joint Commission; URAC = Utilization Review Accreditation Commission.

5.E.7. Weighting

We applied weights to the observations in the regressions to ensure that (1) beneficiaries who were observed for longer periods receive relatively more weight than those observed for shorter periods (using a Medicare enrollment weight), and (2) the CPC+ and comparison groups are comparable (using a matching weight). To achieve the first goal, for each beneficiary in each year, we calculated fractional enrollment weights that capture the share of months observed during that year. For the impact analysis, a beneficiary is observed during each month that he or she is alive and enrolled in Medicare FFS (enrolled in both Part A and Part B, and not in an MA plan), and has Medicare as the primary payer.

As we describe in Appendix 6.C of the appendices to the supplemental volume of the CPC+ evaluation second annual report (Ghosh et al. 2020), we used an external comparison group as the main comparison group for the impact analysis of Medicare claims-based outcomes. For all analyses using this comparison group, the matching weight was the same as the covariate-balancing propensity score-based weights used to balance the CPC+ and comparison practices on their baseline characteristics.

The final composite weight for beneficiaries in the comparison group was the product of (1) the enrollment weight, and (2) the matching weight. For beneficiaries in the CPC+ group, we needed only the enrollment weight because, by construction, the matching weight for each CPC+ beneficiary is one.

Regressions for most outcomes incorporated these final composite weights—that is, the product of the enrollment weight and the matching weight—for CPC+ and comparison beneficiaries in each baseline and intervention period interval. We used slightly different weights for regressions for the following outcomes:

- For discharge-level measures, such as readmissions and unplanned acute care, we incorporated only the matching weight; the enrollment weight was unnecessary, because these regressions included beneficiaries only if they were enrolled in Medicare FFS during the full month following the discharge.[79]
- For certain binary outcomes defined at the beneficiary level—for example, whether a beneficiary received hospice services—we used the composite weight; before doing so, we recoded the enrollment weight to account for truncation due to beneficiaries potentially dying during the follow-up period. Specifically, the enrollment weight was recoded to a value of one if the outcome was observed, to prevent those who received these services from receiving smaller weights due to death, and was equal to the enrollment weight (using the usual methods to take into account length of time observed) if the outcome was not observed.
- For mortality outcomes, such as 12-month mortality in PY 1 and 60-month mortality in PY 5, we used only the matching weights; the enrollment weight was unnecessary because the outcome was observed over a fixed duration of follow-up for all beneficiaries and we know for certain whether a beneficiary was or was not alive at the end of that follow-up period.
- For comprehensiveness of care measures, which were at the practitioner-year level, we used only the matching weight, because there was no weight corresponding to the beneficiary enrollment weight at the practitioner level.

[79] The only exception is that the regression retains beneficiaries who die during the month following the discharge.

5.E.8. Power to detect effects

Given our large sample sizes, the impact analysis is well-powered to detect even small impacts on the primary outcome—Medicare expenditures without CMS's enhanced payments. For both tracks, the power to detect a non-zero effect if the true impact is equal to the average care management fees (CMFs)[80] ($15 per beneficiary per month [PBPM] in Track 1 and $28 PBPM in Track 2) is more than 99 percent over the five program years, and more than 93 percent for each program year. Also, the smallest true effects that the study can detect with at least 80 percent power are $8 and $10 PBPM (less than 1 percent) over the five program years in Track 1 and Track 2, respectively, and between $8 and $14 PBPM (slightly higher than 1 percent) for each program year. Power remains relatively high when we analyze the SSP and non-SSP subgroups separately—for each of the two subgroups, the power to detect non-zero effects on expenditures is at least 92 percent in Track 1 and 99 percent in Track 2 over the five program years, and at least 71 percent in Track 1 and 92 percent for Track 2 for each program year, assuming true effects were equal to the size of the CMFs.

5.E.9. Variation in effects among subgroups of beneficiaries and practices

As we discussed above, within each track, we estimated impacts separately by baseline SSP status of practices to investigate whether participating in both CPC+ and an SSP ACO had a different impact than participating in CPC+ alone. Given that SSP participation is a critical dimension on which participating CPC+ practices differ, we estimated these separate regressions, by SSP status, for all outcomes.

In addition, the impacts of CPC+ could differ for different types of beneficiaries and practices, based on other baseline characteristics. Knowing whether CPC+ is more or less effective for certain types of practices or beneficiaries could inform strategies to help practices succeed. Those findings could also provide insights about the types of practices and beneficiaries who should be encouraged to participate in future primary care transformation efforts like CPC+. Therefore, for selected outcomes, we estimated the effects of the program on *subsets of beneficiaries* for whom CPC+ is likely to have especially large effects, such as the chronically ill and other patients with complex health conditions (Brown et al. 2012; Rich et al. 2012). We also examined effects for different *types of practices,* such as those that had a larger number of primary care practitioners, had participated in prior primary care transformation initiatives at baseline, or were owned by a hospital or health system. For these subgroup analyses, we included in the regressions interactions of variables denoting subgroup membership with the indicator for CPC+ versus comparison status,[81] the intervention year indicator, and the CPC+ indicator interacted with the intervention year indicator. Because there is likely to be significant correlation among practice characteristics, for example, between practice size and ownership, testing for differential effects for each practice characteristic separately may not unmask the real drivers of significant differences. Therefore, for the practice subgroup analysis, we included interactions with subgroup indicators for all practice characteristics in a single regression to disentangle which characteristics actually influence program

[80] Our calculations are conservative in that they assess the power to detect an effect of the size of the CMF; we would have even better power to detect an effect of the size of all of CMS's enhanced payments combined (including the CPC+ CMFs, the comprehensiveness supplement [for Track 2 practices only], Performance-based Incentive Payments, and the payments made to practices' Accountable Care Organizations [ACOs] for SSP shared savings).

[81] The interaction between the practice subgroup membership indicator and the CPC+ indicator cannot be directly estimated in the practice-level subgroup analysis because the model includes practice fixed effects.

impacts.[82, 83] Our main subgroup analyses focus on estimating differential effects for Medicare expenditures without enhanced payments. If we find evidence of differential effects for any particular subgroup(s), we explore it further with additional analyses (for example, by examining effects on service use outcomes for that subgroup, or estimating subgroup effects separately within the SSP and non-SSP samples).

A. Practice-level subgroups

We estimated differential effects for subgroups defined at baseline by various characteristics, as shown in Table 5.E.5.

Table 5.E.5. Practice-level subgroups

Subgroup definitions	Why potentially important to CPC+
Whether the practice had participated in prior primary care transformation initiatives—defined as participation in CPC Classic or the Multi-Payer Advanced Primary Care Practice demonstration, or NCQA, TJC, AAAHC, URAC, or state medical-home recognition status	Practices with participation in prior primary care transformation initiatives may be more advanced and, as a result, may require less time and resources to make changes at the start of CPC+. On the other hand, these practices may have less room for improvement after their prior practice transformation experience.
Practice size, as defined by the number of primary care practitioners (1–2, 3–5, 6 or more)	Larger practices will likely have access to greater resources and better medical infrastructure. Smaller practices may, on the other hand, have greater flexibility to implement changes more rapidly.
Whether the practice was multi-specialty versus primary care only	Multi-specialty practices face different financial incentives and economies of scale.
Whether the practice was owned by a hospital or a health system[a]	Practices owned by a hospital or health system will likely have access to greater resources and better medical infrastructure. These practices may also face different financial incentives and economies of scale.
Whether the practice shared a TIN with another primary care practice	Like practices owned by a hospital or a health system, practices that share a TIN with another primary care practice will likely have access to greater resources and better medical infrastructure. These practices may also face different financial incentives and economies of scale.
Whether the practice was in a rural, suburban, or urban area	Practices in more urban areas will likely have access to greater resources and better medical infrastructure than those in rural areas.

[a] We constructed the variable for hospital or health system ownership at baseline using IQVIA data. We checked this variable against what all responding practices reported in the 2017 practice survey and found good concordance. More than 86 percent of practices that were not hospital- or system-owned according to the IQVIA data reported that they were independent, physician-owned, and less than 7 percent of those classified as owned by a system or hospital in IQVIA data reported that they were independent, physician-owned in the survey.

AAAHC = Accreditation Association for Ambulatory Health Care; NCQA = National Committee for Quality Assurance; TIN = Taxpayer Identification Numbers; TJC = The Joint Commission; URAC = Utilization Review Accreditation Commission.

[82] Given that sharing a TIN is highly correlated with being owned by a hospital or health system, we ran two versions of practice subgroup regressions—one with an indicator of whether practices were owned by a hospital or health system, and the other with an indicator of whether the practice shared a TIN with another primary care practice. All other practice subgroups were included in both versions of regressions.

[83] Given the high degree of overlap between certain beneficiary subgroups—for example, between those above the 75th percentile of the HCC score distribution and those above the 90th percentile—we did not include interactions with all beneficiary subgroup definitions in a single regression. Instead, we estimated a separate regression for each subgroup of interest where we included interactions of treatment (identifying CPC+ practices) and post-intervention (identifying time periods after CPC+ began) indicators with the subgroup indicator denoting whether the beneficiary had that characteristic.

B. Beneficiary-level subgroups

When analyzing differential effects by subsets of beneficiaries, we considered subgroups that tend to have higher utilization and cost, for example, beneficiaries with higher HCC scores or those with behavioral health conditions (Table 5.E.6). As with the beneficiary-level control variables, we identified beneficiary subgroups directly before the start of the baseline period for baseline observations and directly before the start of the intervention period for intervention period observations.

Table 5.E.6. Beneficiary subgroups

Subgroup definitions	Why potentially important to CPC+
Beneficiaries in the highest quartile of the distribution of HCC score (both Track 1 and Track 2), or patients who either were in the highest decile of the distribution of HCC score or had dementia (both Track 1 and Track 2)[a]	Beneficiaries with high HCC scores and/or those with dementia are at greater risk of incurring high health care expenditures. Also, these high-risk definitions are based on CMS's criteria for identifying beneficiaries in risk Tier 4 and risk Tier 5.[b]
Beneficiaries with multiple chronic conditions, specifically at least 2 of 12 frequently occurring chronic conditions,[c] who also had at least one hospitalization in the year before the start of CPC+ (for observations in the intervention period) or the year before baseline (for observations in the baseline period)	Beneficiaries with multiple chronic conditions who have also experienced relatively recent hospitalizations are among the highest-risk beneficiaries.
Beneficiaries who were also eligible for Medicaid (dually eligible)	Dually eligible beneficiaries typically have higher health care utilization and higher costs than those who are not dually eligible.
Beneficiaries with anxiety, depression or substance use disorders	Behavioral health conditions have high prevalence in primary care, frequently co-exist with chronical medical conditions, and are associated with significant overall health care costs.

[a] As with the beneficiary characteristics, the HCC score or conditions used to define these subgroups are measured directly before the start of CPC+ (for the intervention-period observations) or directly before the start of the yearlong baseline period (for the baseline-period observations). We exclude new enrollees from these subgroup analyses since their HCC scores and HCCs are based on demographic characteristics only and we cannot reliably assess their actual risk status in the absence of claims data.

[b] CMS's approach for identifying Tier 4 and Tier 5 high-risk beneficiaries differs from the approach we used in the impact analysis. Specifically, CMS includes the entire Medicare population in each CPC+ region, and uses the region-specific distribution of HCC scores to identify the 75th and 90th percentiles of the distribution. For the impact analysis, we identified the high-risk HCC cutoffs by looking at the distribution of 2016 HCC scores among Medicare beneficiaries in our baseline sample, and across all regions. Also, CMS identifies Tier 5 patients for Track 2 only, whereas we also ran subgroup analyses for Tier 5 beneficiaries in Track 1 practices. Details of our methodology for calculating HCC scores and how it deviated from CMS's approach are in Appendix 5.B, Section 5.E.3.

[c] The 12 frequently occurring chronic conditions we used in this definition are: congestive heart failure, chronic obstructive pulmonary disease, acute myocardial infarction, ischemic heart disease, diabetes, metastatic cancer and acute leukemia, stroke, depression, dementia, atrial fibrillation, rheumatoid arthritis or osteoarthritis, and chronic kidney disease. These chronic conditions are measured by HCCs (or combinations of HCCs) except for dementia, which is measured using the indicator for Alzheimer's disease or dementia from the Chronic Conditions Warehouse, and chronic kidney disease, which is measured using the original reason for entitlement to Medicare being ESRD.

CMS = Centers for Medicare & Medicaid Services; ESRD= end-stage renal disease; HCC = hierarchical condition category.

For all subgroup analyses, we checked the percentage of the CPC+ and comparison groups that belonged to each subgroup category to ensure similarity in the percentages across the two groups. We also examined key baseline characteristics we used in matching, such as Medicare expenditures, acute care hospitalizations, and outpatient ED visits to check the similarity of the CPC+ and comparison groups

within each subgroup. For most characteristics, CPC+ and comparison groups were well-balanced within each subgroup.[84, 85] This was also true for key baseline characteristics within subpopulations used in examining specific outcomes, such as beneficiaries ages 18 through 75 (the subpopulation used for the diabetes measure), female beneficiaries ages 52 through 74 (the subpopulation used for the breast cancer screening measure), beneficiaries with a minimum number of ambulatory care visits (the subpopulation used for the continuity-of-care measures), beneficiaries who are continuously enrolled in Medicare Part D and have a relevant diagnosis (the subpopulation used for outcomes related to appropriate use of medications), and beneficiaries continuously enrolled in Medicare Part A and B FFS for the measurement year and the preceding year (the subpopulation for low-value services measure).

C. Checking for differences in impact estimates by subgroup

The following steps describe the process we used to check for differences in impact estimates by practice subgroup:

1. To test for significant differences across all subgroups defined by practice characteristics, we conducted a joint test of significance across all subgroups to determine whether there was any evidence of variation in impacts across practice subgroups in general. This approach helped minimize the number of tests checking for statistically significant differences across subgroups and reduced the likelihood of erroneously concluding that a chance difference across subgroups was meaningful. If we were unable to reject the null hypothesis in this test of no difference across the range of subgroups defined by all practice characteristics, we considered any evidence of differences across subgroups defined by a *single* characteristic to be weak.
2. For subgroups defined by any particular practice characteristic, we tested whether the impact estimates for the subgroups defined by the same characteristic were significantly different from one another:[86]
   - **a.** If this test did not show a statistically significant difference, we concluded that there was no meaningful difference in impact estimates for subgroups defined by that particular practice characteristic.
   - **b.** Only if this test showed a statistically significant difference ($p < 0.10$) did we test for whether the impact estimate *within* the subgroup was significantly different from zero.

[84] We considered CPC+ and comparison groups to be well-balanced on a characteristic if the standardized difference was 0.25 standard deviations or less. We found that, for most characteristics, the standardized difference was well under the 0.25 threshold within each subgroup.

[85] The only subgroups where the standardized differences were higher than 0.25 for more than 10 percent of the variables (out of 61 variables examined in total) were practices located in rural or suburban counties and practices that did not share a TIN with another primary care practice. For these subgroups, the higher standardized differences were mostly found among lower-priority variables (such as region indicators, county-level poverty rates, the number of hospital beds, or the median household income in the county).

[86] We conducted the test for statistically significant differences across subgroups defined by a single characteristic, even if the null hypothesis in the joint significance test was not rejected—that is, even if the evidence for variation in impact estimates across subgroups was weak from the joint test of significance across all subgroups. If the joint test across all subgroups was not statistically significant, we would more cautiously interpret any statistically significant difference between subgroups defined by a single characteristic.

For example, for the subgroup defined by prior experience with primary care transformation, we first tested whether the impact estimates for practices that participated in prior transformation activities and those that did not were significantly different from one another. If the p-value from this test did not lead us to reject the hypothesis that the impacts were similar, we concluded that impacts did not vary meaningfully across subgroups defined by prior experience with primary care transformation. On the other hand, if this test showed a statistically significant difference ($p < 0.10$), we then tested whether the impact estimate within each subgroup—practices that participated in prior transformation activities and those that did not—was significantly different from zero.

As noted above, for subgroups defined by beneficiary characteristics, we estimated a separate regression for each subgroup of interest. Consequently, we did only Step 2 of the above process for beneficiary subgroup analyses.

5.E.10. Sensitivity tests

We calculated alternative estimates as robustness checks of the main impact estimates on the key outcomes of Medicare expenditures, acute hospitalizations, and outpatient ED visits. Specifically, we assessed the sensitivity of our results to changes in the following key elements of our estimation approach: (1) definition of the beneficiary sample, (2) modeling assumptions, (3) length of the baseline period, (4) controlling for contemporaneous (same year) SSP participation status, and (5) alternative definition of the counterfactual (by using a triple-differences approach). We also conducted COVID-19-specific sensitivity tests by examining impact estimates after excluding claims from the peak COVID-19 period (March–May 2020). We also conducted a sensitivity test for readmissions and unplanned acute care outcomes by defining the outcome at the beneficiary level instead of at the discharge level. We describe the motivation for each sensitivity test in Table 5.E.7.

When results from the sensitivity tests were inconsistent with results from our main analysis, we incorporated that information into our discussion and interpretation of findings. We assessed the conditions under which the alternative estimates would be preferred, and the likelihood that those conditions were met.

Table 5.E.7. Sensitivity tests

Sensitivity test	Motivation
Altering the composition of the beneficiary sample	
Use sample of beneficiaries attributed during the intervention period (who are also attributed during the baseline period) as the baseline sample	Helps to adjust for changes in sample composition between baseline and follow-up that may differ for the intervention and comparison groups
Examine impacts for the subset of Medicare beneficiaries attributed in the first quarter of the period (that is, the first quarter of the baseline period and the first quarter of the intervention period)	Removes effects that may be due to differences over time in sample additions between the intervention and comparison groups. This might occur if, for example: (1) different types of beneficiaries are attracted to receive care at CPC+ practices than at comparison practices, (2) CPC+ and comparison practitioners have incentives to retain or dismiss certain types of patients, or (3) a higher proportion of beneficiaries are attributed to the CPC+ than comparison practices over time via Annual Wellness Visits
Instead of following an intent-to-treat (ITT) approach to defining the beneficiary sample (once attributed, beneficiaries stay in the sample for the rest of the baseline or intervention period), allow beneficiaries to drop out of the sample, if they no longer meet attribution requirements	Assesses whether the ITT approach tends to attenuate true effects by retaining beneficiaries in the intervention group who are no longer seen by CPC+ practices
Altering the modeling assumptions	
For analysis of expenditures, use a generalized linear model with log link	Accounts for skewed expenditure distribution
Log-transform the expenditures variable (generating impact estimates in percentage terms)	Reduces influence of high-cost cases; accounts for skewed expenditure distribution
Trim expenditures at 98th percentile	Reduces influence of high-cost cases
Use baseline beneficiary characteristics, practice characteristics, and practice-level averages of beneficiary characteristics (reflecting baseline characteristics of contemporaneous beneficiaries), all interacted with year indicators as additional controls	Accounts for potential time-varying effects of baseline beneficiary and practice characteristics on the outcome. Adjusts for practice-level measures of beneficiary characteristics to align with participation in CPC+ varying at the practice level
Altering length of baseline period	
Use two instead of one pre-intervention years in the baseline period	Tests whether impact estimates are sensitive to using a longer baseline period and whether there are differences in trends prior to CPC+ for CPC+ and comparison practices
Controlling for contemporaneous SSP participation	
Use contemporaneous (same year) SSP status instead of baseline SSP status as a covariate or to separately examine impacts for SSP/non-SSP subgroup	Accounts for any difference in contemporaneous SSP participation between the CPC+ and comparison groups and its effect on outcomes
Alternative definition of counterfactual	
Use a triple-differences model and include non-participating practices in CPC+ regions and unselected practices in comparison regions in the analytic sample	Accounts for regional shocks that might affect CPC+ and comparison regions differently (see Section 5.E.11 for details)
COVID-19-specific sensitivity tests (for PY 4 estimate)	
Examine impacts after excluding the first three months of COVID-19 (March–May 2020)	Tests for the sensitivity of the estimate to the reduction in service utilization in the peak COVID-19 period

Table 5.E.7 ***(continued)***

Sensitivity test	Motivation
Definition of outcome measures	
Examine impacts on the beneficiary-level readmission and unplanned acute care outcomes, defined as the probability of readmission or unplanned acute care after an index discharge[a] during a year	Removes concerns about possible endogeneity in analysis of readmission and unplanned acute care outcomes, which can arise if CPC+ alters the probability of an index discharge. In that case, the analysis of the discharge-level readmission and unplanned acute care measures would be biased, because CPC+ may have prevented hospitalizations or ED visits or observation stays that would have been at lower relative risk of a readmission or receipt of unplanned acute care
Use expenditures that exclude the QPP payments	Tests whether estimates are sensitive to an alternative definition of the primary outcome measures – Medicare expenditures without enhanced payments

[a] An index discharge refers to an index hospital discharge (for the outcomes of readmission and unplanned acute care after hospitalization) or an index ED visit or observation stay (for the outcome of unplanned acute care after an ED visit or an observation stay).

ED = emergency department; ITT = intent-to-treat; PY = Program Year; QPP = Quality Payment Program; SSP = Medicare Shared Savings Program.

5.E.11. Triple-differences analysis

In this section, we describe the triple-differences model used to examine the sensitivity of the main impact estimates to potential bias due to inadequately accounting for differences in regional trends. Because the comparison practices are from non-CPC+ regions, they may experience different trends in outcomes (potentially due to different market conditions or regional shocks) than CPC+ practices do, which may cause our impact estimates to reflect these differential regional trends rather than the causal impacts of CPC+ itself. We first explain the study population, unit of observation, and outcomes in the triple-differences regressions (Section 5.E.11.A). We then present details on model specification and estimation (Sections 5.E.11.B to 5.E.11.D), followed by a description of weighting (Section 5.E.11.E) and the joint significance test of the differences between the difference-in-differences estimates and the triple-differences estimates (Section 5.E.11.F). Finally, we discuss the sensitivity tests to check for the robustness of the triple-differences estimates (Section 5.E.11.G).

A. Study population, unit of observation, and outcomes

Sample of practices. Our sample includes practices in regions that include the 2017 CPC+ Starters and practices in the comparison regions. The set of practices includes CPC+ and comparison practices, as well as non-CPC+ practices and non-comparison practices, which are primary care practices in the same regions as CPC+ and comparison practices that did not participate in CPC+ or were not selected as comparison practices. For non-CPC+ practices and non-comparison practices, we applied the same practice exclusion criteria used in selecting the comparison group described in Appendix 6.C of the appendices to the supplemental volume of the CPC+ evaluation second annual report (Ghosh et al. 2020).

Beneficiary assignment based on attribution. To estimate the triple-differences model, we used an intent-to-treat (ITT) analysis approach that includes practices described above and their "assigned" beneficiaries. Our beneficiary assignment was consistent with the approach for the main impact analysis. That is, once we attributed a beneficiary to a CPC+ or comparison practice in any baseline or intervention quarter, we continued to assign that beneficiary to the same practice in future baseline and intervention quarters, regardless of whether the beneficiary continued to receive care at that practice. However, if a beneficiary was at first attributed to a non-CPC+ practice or a non-comparison practice during the

intervention period, but later attributed to a CPC+ or comparison practice in subsequent program years, that beneficiary would be re-assigned to that CPC+ or comparison practice in the subsequent program years. We did this to ensure similarity between the difference-in-differences and triple differences CPC+ and comparison samples.

Table 5.E.8 shows the number of practices and the number of Medicare FFS beneficiaries in the triple-differences analysis and in the main impact analysis, for each track and practice group. In previous reports, the triple-differences sample contained the same number of CPC+ and comparison practices, but a slightly higher number of unique beneficiaries assigned to these practices (less than 1 percent higher) for both Track 1 and Track 2, due to minor adjustments to the ITT approach compared to that used in the main analysis.[87] In the current analysis, we dropped these extra beneficiaries, so the samples used in the triple-differences analysis and the main analysis are the same.

Table 5.E.8. Numbers of practices and of Medicare FFS beneficiaries in the triple-differences analysis and the difference-in-differences analysis, by track and practice group

	CPC+		Comparison		Non-CPC+		Non-comparison	
Research sample	Triple-differences	Difference-in-differences	Triple-differences	Difference-in-differences	Triple-differences	Difference-in-differences	Triple-differences	Difference-in-differences
Track 1								
Number of practices	1,373	1,373	5,243	5,243	8,337	n.a.	20,656	n.a.
Number of beneficiaries	1,549,585	1,549,585	5,347,499	5,347,499	4,015,775	n.a.	11,444,943	n.a.
Track 2								
Number of practices	1,515	1,515	3,783	3,783	7,276	n.a.	20,115	n.a.
Number of beneficiaries	1,896,880	1,896,880	4,507,499	4,507,499	3,378,353	n.a.	11,153,265	n.a.

Source: Mathematica's analysis of Medicare claims data from January 2014 through December 2021.

FFS = fee-for-service; n.a. = not applicable; non-comparison = unselected practices in comparison regions; non-CPC+ = non-participating practices in CPC+ regions.

Unit of observation. The unit of observation in the regressions is the beneficiary-year. Each beneficiary has observations for as many years as the person remains in the sample and can still be observed in Medicare claims. The observability criteria are the same as in the main impact analysis. Specifically, to be observed, a beneficiary assigned to a practice for the baseline or the intervention period had to be alive, have both Part A and B Medicare FFS coverage with Medicare as the primary payer, and not be covered under a Medicare Advantage or other Medicare health plan.

[87] Specifically, for the triple-differences analysis, we allowed (1) 2018 Starter comparison practices in 2017 Starter comparison regions to be non-comparison practices, (2) practices that applied to CPC+ but were not selected to participate to be non-CPC+ practices, and (3) the baseline and intervention periods for the non-CPC+ and non-comparison practices to be the same as those for the 2017 Starters. For example, we expected allowing 2018 Starter comparison practices in 2017 Starter comparison regions to be non-comparison practices to increase the number of beneficiaries assigned to 2017 CPC+ or comparison practices, because beneficiaries attributed to 2018 Starter comparison practices in PY 1 could switch into 2017 Starter CPC+ or comparison practices in later program years.

Outcomes. We defined a set of main outcomes that represent key hypothesized effects of the model for which all subgroup analyses and sensitivity tests were conducted. We also identified a set of secondary outcomes which were particularly impacted by COVID-19, or that provide additional context to the results for the main outcomes.

- Main outcomes:
  - Medicare Part A and B expenditures without enhanced payments for CPC+ and SSP, in dollars per beneficiary per month
  - Annualized number of acute hospitalizations per 1,000 beneficiaries
  - Annualized number of outpatient ED visits per 1,000 beneficiaries
- Secondary outcomes:
  - Annualized ambulatory primary care visits per 1,000 beneficiaries
  - Annualized urgent care center visits per 1,000 beneficiaries
  - Annualized non-face-to-face primary care visits as a portion of all ambulatory primary care visits per 1,000 beneficiaries (telehealth)

We conducted analyses by practices' SSP participation for all outcomes. We only conducted sensitivity analyses for the triple-differences model for the Medicare expenditures outcome.

B. Model specification

Main model. For all outcomes except the proportion of primary care ambulatory visits that were non-face-to-face, we used the following specification:

Let *i* index the beneficiary; *j* index the practice; and *t* index time, where *t* ranges from 0 to 5, with 0 denoting the baseline year. We estimated a triple-differences regression model for beneficiaries assigned to CPC+ practices, selected comparison practices, non-CPC+ practices, and non-comparison practices. The model had the following form:

$$(5.E.3)\quad y_{ijt} = \alpha + \beta X_{it} + \pi X_{it} s_j + \gamma_t p_t + \theta_t a_j p_t + \varphi_t s_j p_t + \delta_t C_{jt} p_{t\{t=2020,2021\}} + \mu_t a_j s_j p_t + b_j + \varepsilon_{ijt},$$

where

y_{ijt} is an outcome variable for beneficiary *i*, in practice *j*, in year *t*.

X_{it} is a vector of beneficiary characteristics, consistent with those used in the difference-in-differences model (Section 5.E.6.A).

p_t (for "post") is an intervention-period indicator that takes the value of 1 during a specific program year, for instance PY 1, and 0 otherwise.

C_{jt} is a vector of COVID-19-related controls including excess deaths in the state-hospital referral region (HRR), Pandemic Vulnerability Index in the county, Government Response Index in the state, and Social Vulnerability Index in the census tract of each practice. We included COVID-19-related controls interacted with the contemporaneous year indicator (2020 or 2021) to account for potential effects of COVID-19 on outcomes in calendar years 2020 and 2021, respectively.[88]

b_j is a practice-level fixed effect for practice *j*, which controls for all time-invariant practice characteristics.

a_j (for "area") is a binary indicator for being in a CPC+ region; the indicator takes the value of 1 if the practice *j* is located in a CPC+ region and is 0 otherwise. The main effect of this indicator is not identified in this equation since it is collinear with the practice fixed effects.

s_j (for "selected") is a binary indicator for being a CPC+ or comparison practice; the indicator takes the value of 1 if practice *j* is a CPC+ practice or a comparison practice, and is 0 if practice *j* is a non-CPC+ practice or a non-comparison practice. The main effect of this indicator is not identified in this equation since it is collinear with the practice fixed effects.

ε_{ijt} is the idiosyncratic error term. It represents unexplained variability in the outcome variable for beneficiary *i*, in practice *j*, during year *t*.

Our coefficients of interest are the μ_t, which represent the triple-differences impact for each of the five program years. Table 5.E.9 summarizes how we used the parameter estimates from Equation (5.E.3) to obtain the regression-adjusted group means for CPC+ practices, comparison practices, non-CPC+ practices, and non-comparison practices, for the baseline and five program years.

[88] The main effects of these COVID-19-related regional controls are not identified in this equation because the model includes practice fixed effects.

Table 5.E.9. Impact estimate and group means for CPC+ practices, comparison practices, non-CPC+ practices, and non-comparison practices based on a linear regression from Equation (5.E.3)

Comparison regions

Year	Comparison group mean	Non-comparison group mean	Difference between comparison and non-comparison group means	Difference-in-differences	Triple differences
Baseline year (t=0) [reference period]	$\alpha+\pi+(\sigma)$	α	$\pi+(\sigma)$	(φ)	N/A
PY 1 (t=1)	$\alpha+\pi+\gamma_1+\delta_1+(\sigma)$	$\alpha+\gamma_1$	$\delta_1+\pi+(\sigma)$	$(\varphi)+\theta_1$	N/A
PY 2 (t=2)	$\alpha+\pi+\gamma_2+\delta_2+(\sigma)$	$\alpha+\gamma_2$	$\delta_2+\pi+(\sigma)$	$(\varphi)+\theta_2$	N/A
PY 3 (t=3)	$\alpha+\pi+\gamma_3+\delta_3+(\sigma)$	$\alpha+\gamma_3$	$\delta_3+\pi+(\sigma)$	$(\varphi)+\theta_3$	N/A
PY 4 (t=4)	$\alpha+\pi+\gamma_4+\delta_4+(\sigma)$	$\alpha+\gamma_4$	$\delta_4+\pi+(\sigma)$	$(\varphi)+\theta_4$	N/A
PY 5 (t=5)	$\alpha+\pi+\gamma_5+\delta_5+(\sigma)$	$\alpha+\gamma_5$	$\delta_5+\pi+(\sigma)$	$(\varphi)+\theta_5$	N/A

CPC+ regions

Year	CPC+ group mean	Non-CPC+ group mean	Difference between CPC+ and non-CPC+ group means	Difference-in-differences	Triple differences
Baseline year (t=0) [reference period]	$\alpha+\pi+(\rho+\sigma+\tau)$	$\alpha+(\rho)$	$\pi+(\sigma+\tau)$	(φ)	N/A
PY 1 (t=1)	$\alpha+\pi+\gamma_1+\theta_1+\delta_1+\mu_1+(\rho+\sigma+\tau)$	$\alpha+\gamma_1+\theta_1+(\rho)$	$\pi+\delta_1+\mu_1+(\sigma+\tau)$	$\delta_1+\mu_1$	μ_1
PY 2 (t=2)	$\alpha+\pi+\gamma_2+\theta_2+\delta_2+\mu_2+(\rho+\sigma+\tau)$	$\alpha+\gamma_2+\theta_2+(\rho)$	$\pi+\delta_2+\mu_2+(\sigma+\tau)$	$\delta_2+\mu_2$	μ_2
PY 3 (t=3)	$\alpha+\pi+\gamma_3+\theta_3+\delta_3+\mu_3+(\rho+\sigma+\tau)$	$\alpha+\gamma_3+\theta_3+(\rho)$	$\pi+\delta_3+\mu_3+(\sigma+\tau)$	$\delta_3+\mu_3$	μ_3
PY 4 (t=4)	$\alpha+\pi+\gamma_4+\theta_4+\delta_4+\mu_4+(\rho+\sigma+\tau)$	$\alpha+\gamma_4+\theta_4+(\rho)$	$\pi+\delta_4+\mu_4+(\sigma+\tau)$	$\delta_4+\mu_4$	μ_4
PY 5 (t=5)	$\alpha+\pi+\gamma_5+\theta_5+\delta_5+\mu_5+(\rho+\sigma+\tau)$	$\alpha+\gamma_5+\theta_5+(\rho)$	$\pi+\delta_5+\mu_5+(\sigma+\tau)$	$\delta_5+\mu_5$	μ_5

Notes: To highlight the key coefficients in Equation (5.E.3) above, we exclude the coefficients on beneficiary characteristics, COVID-19-related controls, and practice fixed-effects in the expressions for group means in this table. The parameter ρ denotes a coefficient on the indicator for being in a CPC+ region, the parameter σ denotes a coefficient on the indicator for being a CPC+ or comparison practice, and the parameter τ denotes a coefficient on the interaction between the indicator for being in a CPC+ region and the indicator for being a CPC+ or comparison practice. ρ, σ, and, τ, are not included in Equation (5.E.3); they cannot be directly estimated because the model includes practice fixed effects. We include these terms in this table to illustrate the difference-in-difference-in-differences approach, but we show it in parentheses since we did not obtain the estimates. These parameters are differenced out in obtaining the impact estimate.

Non-comparison = unselected practices in comparison regions; non-CPC+ = non-participating practices in CPC+ regions; PY = Program Year.

Model for the proportion of ambulatory primary care visits that are not face-to-face. Since non-face-to-face ambulatory primary care visits were essentially zero for both CPC+ and comparison practices prior to the COVID-19 pandemic in PY 4 (2020), we used an alternative specification for this outcome that did not use data prior to PY 4. In this model, we take the difference in the average outcome between CPC+ and non-CPC+ practices in PY 4 or PY 5, and subtract from that the difference in the average outcome between comparison and non-comparison practices in PY 4 or PY 5. Similar to the main impact analysis for this outcome, we estimate the difference between the two program years separately. The triple-differences model (which would then subtract the same quantity in the baseline year) would yield similar results since the quantity subtracted would be close to 0. The advantage of this model over a triple-differences specification is that it is more transparent about how the estimate is being identified. Our approach mirrors how we estimate the impacts of CPC+ on non-face-to-face ambulatory primary care visits in our main difference-in-differences analysis: we only compare outcomes among CPC+ practices to comparison practices in 2020 and 2021, and not in prior years.

Specifically, we used the following specification. Let *i* index the beneficiary, and *j* index the practice. We estimated a modified difference-in-differences regression model for beneficiaries assigned to CPC+ practices, selected comparison practices, non-CPC+ practices, and non-comparison practices during PY 4 or PY 5 only. The model had the following form:

$$(5.\text{E}.4)\quad y_{ij} = \alpha + \beta X_i + \gamma D_j + \delta s_j + \pi X_i s_j + \rho C_j + \varphi D_j s_j + \theta a_j + \mu a_j s_j + \varepsilon_{ij},$$

where

y_{ij} represents the proportion of ambulatory primary care visits that were not face-to-face in PY 4 or PY 5, for beneficiary *i* in practice *j*.

X_i is a vector of characteristics of beneficiary *i* that includes those from Equation (5.E.3) as well as baseline Medicare expenditures and use of selected services to account for differences in health care utilization between beneficiaries assigned to CPC+ versus comparison practices before the start of CPC+. We describe the baseline Medicare expenditures and service use control variables in more detail in Section 5.E.6 above.

C_j is a vector of COVID-19-related controls as in Equation (5.E.3). We included COVID-19-related controls for examining telehealth outcomes in PY 4 or PY 5.

D_j is a vector of characteristics of practice *j* measured at baseline. We describe practice-level control variables in more detail below.

s_j (for "selected") is a binary indicator for being a CPC+ or comparison practice; the indicator takes the value of 1 if the practice *j* is a CPC+ practice or a comparison practice, and is 0 if practice *j* is a non-CPC+ practice or a non-comparison practice.

a_j (for "area") is a binary indicator for being in a CPC+ region; the indicator takes the value of 1 if the practice j is located in a CPC+ region and is 0 otherwise.

ε_{ij} is the idiosyncratic error term. It represents unexplained variability in the outcome variable for beneficiary i and in practice j.

μ is the modified difference-in-differences impact in PY 4 or PY 5. Table 5.E.10 summarizes how we used the parameter estimates from Equation (5.E.4) to obtain the regression-adjusted group means for CPC+ practices, comparison practices, non-CPC+ practices, and non-comparison practices in PY 4 or PY 5.

Table 5.E.10. Impact estimate and group means for CPC+ practices, comparison practices, non-CPC+ practices, and non-comparison practices for the non-face-to-face ambulatory primary care visit outcome in PY 4 or PY 5 based on a linear regression from Equation (5.E.4)

Comparison regions

Comparison group mean	Non-comparison group mean	Difference between comparison and non-comparison group means	Difference-in-differences
$\alpha+\pi+\varphi+\delta$	α	$\pi+\varphi+\delta$	N/A

CPC+ regions

CPC+ group mean	Non-CPC+ group mean	Difference between CPC+ and non-CPC+ group means	Difference-in-differences
$\alpha+\pi+\varphi+\theta+\delta+\mu$	α	$\pi+\varphi+\delta+\mu$	μ

Notes: To highlight the key coefficients in Equation (5.E.4) above, we exclude the coefficients on beneficiary characteristics, COVID-19-related controls, and practice characteristics in the expressions for group means in this table.

Non-comparison = unselected practices in comparison regions; non-CPC+ = non-participating practices in CPC+ regions.

C. Control variables

Main model controls. We included the same set of beneficiary characteristics as in the main impact analysis (see Table 5.E.4 for the list of beneficiary-level controls). To allow for the possibility that beneficiary characteristics might have different effects for beneficiaries in CPC+ or comparison practices than for beneficiaries in non-CPC+ or non-comparison practices, we interacted the beneficiary control variables with an indicator for whether the beneficiary was assigned to a CPC+ or comparison practice.

Consistent with the main impact analysis, we included the same set of covariates that capture the magnitude of the pandemic within the state as well as the strength of state-level policy responses, including excess deaths in the state-HRR, the Pandemic Vulnerability Index in the county, the Government Response Index in the state, and the Social Vulnerability Index in the census tract. Each year-specific COVID-19-related variable is interacted with the contemporaneous year indicator (2020 or 2021).[89] We also interacted them with an indicator for whether a beneficiary was assigned to a practice included in the evaluation (that is a CPC+ or comparison practice), to allow for the possibility that COVID-19-related variables might have different effects for beneficiaries in CPC+ or comparison practices than for beneficiaries in non-CPC+ or non-comparison practices.

Non-face-to-face ambulatory primary care visits model controls. Because the model for the proportion of ambulatory primary care visits that were not face-to-face compares CPC+, comparison, non-CPC+, and non-comparison practices only in PY 4 and 5, we could not include practice fixed-effects as the other models do.[90] Similar to the difference-in-differences model for telehealth outcomes in the main analysis, we instead included a vector of detailed practice and region characteristics (see Table 5.E.4). Also, similar to the main analysis, to adjust for differences in health care utilization among beneficiaries attributed to the practice at baseline, we included the average monthly Medicare expenditures, annualized number of acute hospitalizations, outpatient ED visits, and ambulatory primary care visits at baseline, and an indicator for whether baseline Medicare expenditures and service utilization were missing at the practice level. We also interacted each of these variables with an indicator for whether the beneficiary was assigned to a CPC+ or comparison practice.

D. Model estimation

Our model estimation approach was the same as in the main impact analysis:

- The regression sample included the baseline year (2016) and the five intervention years (PY 1, PY 2, PY 3, PY 4 and PY 5) for all outcomes except the proportion of ambulatory primary care visits that were not face-to-face, which only includes the fourth and fifth intervention years (PY 4 and PY 5).
- We estimated Equations (5.E.3) and (5.E.4) as linear regressions, separately for Track 1 and Track 2, and also separately by SSP status within each track.

[89] In prior years, we did not include COVID-19 controls in the main triple-differences analysis, but added them to our models in a sensitivity test of the triple-differences model's key assumption that non-CPC+ and non-comparison regions experienced impacts of COVID-19 that were comparable to their CPC+ and comparison region counterparts. This year, we included the COVID-19 controls in the main triple-differences analysis in order to "nest" the main impact model within the triple-differences model and directly compare findings.

[90] Practice fixed effects capture time-invariant variation in practice characteristics and are therefore appropriate in models measuring practices' outcomes across multiple years. In the non-face-to-face ambulatory primary care visits model, we included only one year of data (PY 4 or PY 5). Including practice fixed effects would therefore eliminate the variation in CPC+, non-CPC+, comparison, and non-comparison practices' outcomes from which the impact estimate is derived.

- All regressions accounted for non-independence across observations within the same practice, using standard error estimates clustered at the practice level.
- Each regression included practice fixed effects, except the model for proportion of ambulatory primary care visits that were not face-to-face, where practice-level controls were included.

E. Weighting

For beneficiaries in CPC+ or comparison practices, we applied the same weights as in the main impact analysis. That is, the final weight for beneficiaries in the comparison group was the product of the enrollment weight and the matching weight. For beneficiaries in the CPC+ group, we needed only the enrollment weight, because, by construction, the matching weight for each CPC+ beneficiary is 1.

For beneficiaries in non-CPC+ practices or non-comparison practices, the final weight was the product of the enrollment weight and the baseline concentration weight. We constructed the concentration weight at the state-HRR level such that non-CPC+ practices had the same level of representation (in terms of beneficiary months in the baseline year) as CPC+ practices of the same SSP-status in the same state and HRR, and likewise that non-comparison practices had the same level of representation as weighted comparison practices of the same SSP-status in the same state and HRR.[91,92]

F. Testing joint significance of differences between main impact analysis and triple-differences analysis

We report the statistical significance of the *difference* between individual impact estimates from the main impact analysis versus the triple-differences analysis, captured by δ_t. Testing for broader patterns in differences between the main impact analysis and the triple-differences analysis provides two benefits. First, it allows us to identify whether differences in the two models' predictions are consistent across outcomes in a given year, potentially reflecting the effects of different modeling assumptions during specific timeframes such as the peak of the pandemic. Second, across six primary and secondary outcomes and five program years, individual tests of differences between the two models suffer from multiple comparison issues that may result in false positives. Joint tests of these differences are better able to identify divergences between the two models that are consistent across outcomes and years. To identify these broader patterns, we tested for the joint significance of δ_t across outcomes and/or program years in the following ways, using a Seemingly Unrelated Regressions model (we did not test for joint significance of differences across cumulative impact estimates):

- Equality of annual impact estimates for priority outcomes (Medicare expenditures without enhanced payments, acute hospitalizations, and outpatient ED visits) between the main impact analysis and the triple-differences analysis, for each year.

[91] The only exception to the balanced representation at the state-HRR level is for state-HRRs that had only CPC+ or comparison practices of a given SSP status, in which case there is no representation of non-participating practices or unselected practices of the same SSP status in those specific state-HRRs. We adjusted the concentration weight for practices that are in the same state for such cases so that the representation at state level was still balanced.

[92] We updated the concentration weight for the triple-differences analysis in the fourth annual report to ensure balance by SSP status and also to ensure that non-comparison practices had the same level of representation as *matching weighted* comparison practices.

- Equality of annual impact estimates for priority outcomes (Medicare expenditures without enhanced payments, acute hospitalizations, and outpatient ED visits) between the main impact analysis and the triple-differences analysis, across all program years.

G. Sensitivity analyses

We conducted the following sensitivity tests to assess the robustness of the findings from the triple-differences analysis:

- Winsorize the concentration weight at the 99th percentile. This test helped to check if extreme values of the concentration weight could be driving the findings.
- Alternate model specification omitting control variables related to COVID-19 and their interactions with beneficiary treatment status. In prior years, these controls were added to test whether non-CPC+ and non-comparison practices' outcomes accurately reflected the impact of COVID-19 on CPC+ and comparison practices' outcomes. This year, these variables are included in the main specification. If including these variables has additional explanatory power in the triple-differences model (and omitting these variables therefore causes our impact estimates to change), this would suggest that non-CPC+ and non-comparison practices' outcomes do not fully capture the effects of COVID-19 on CPC+ and comparison practices.
- Not use the concentration weight for non-CPC+ and non-comparison practices. If the number of practices (and their beneficiaries) changes differentially across the analysis groups during the intervention period (for example, due to differences in practice closures or COVID-19 related mortality), the *baseline* concentration weight may no longer lead to similar levels of geographic representation between analysis groups during the *intervention* period. As a result, the triple-differences model would not cancel out the regional shocks as intended. This check helps to assess if the findings are sensitive to the use of concentration weights.
- Exclude non-CPC+ practices (and non-comparison practices) that had the same Taxpayer Identification Number (TIN) as CPC+ (or comparison) practices.[93] This test helps to check if the triple-differences estimates are robust to the potential spillover of any favorable impact of CPC+ to non-participating practices owned by the same parent entity. If there are favorable spillovers, we would be netting out part of the effect of CPC+ in the triple-differences analysis, which would dilute the estimated effects of the intervention.
- Include only beneficiaries attributed in the first quarter of baseline and intervention period. This test checked whether the triple-differences estimates may be driven by differential trends in patient migration into and out of practices. If beneficiaries newly attributed to CPC+, comparison, non-CPC+, and non-comparison practices differ systematically over the intervention period, our impact estimates may reflect the changing composition of attributed beneficiaries rather than a causal impact of CPC+.[94]

[93] Because excluding the TIN-sharing non-CPC+ practices changes the composition of practices in the non-CPC+ sample, we excluded non-comparison practices that share TINs with comparison practices to make the remaining sample of non-CPC+ and non-comparison practices more comparable.

[94] This may be particularly a concern for the triple-differences analysis because of the change in practice rosters from SK&A to OneKey in 2019. While we tracked CPC+ and comparison practices as closely as possible over the data transition, we did not do the same for other practices, including the non-CPC+ and non-comparison practices that were included in the triple-differences sample.

Although the triple-differences model is more robust to the presence of differential regional shocks or trends between CPC+ and comparison regions, the estimates from this model should still be interpreted in the context of the triple-differences model's limitations. For example, the triple-differences model nets out any potential positive spillovers (for example, knowledge of practice transformation) that could flow from CPC+ to non-CPC+ practices within the region, and thus omits a portion of CPC+'s potential effects. Second, the triple-differences model estimates are generally less precise, which makes it less likely that the model would accurately detect small, yet policy-relevant, program impacts. Third, although the CPC+ and comparison practices were matched on all key practice characteristics and outcomes in the baseline period using propensity score matching methods, the non-CPC+ and non-comparison practices are not as well matched on practice characteristics. If regional shocks have differential effects by practice type, then the regional trend that we are netting out by using the non-treatment and non-comparison practices may not be the right counterfactual for the regional trend that would be experienced by CPC+ or comparison practices. Fourth, when estimating the impacts of CPC+ across five program years, two tracks, and two SSP subgroups, the number of individual impact estimates becomes large enough that statistically significant impact estimates may occur purely by chance, leading to disagreements between the difference-in-differences and triple-differences models' estimates.

5.F. Participation in other initiatives

CPC+ is taking place at the same time as many other initiatives that aim to improve the quality and value of medical care. CPC+ practices are allowed to participate in some, but not all, of these initiatives; therefore, we expect comparison practices to participate in some initiatives—such as billing for chronic care management (CCM) services—at higher rates than the CPC+ practices. Higher participation rates among comparison practices than among CPC+ practices will not bias our main impact estimates, because we assume that the comparison practices represent the accurate counterfactual for CPC+ practices had CPC+ not existed (that is, CPC+ practices might have participated in other initiatives at higher rates had CPC+ not existed). At the same time, differences in participation could potentially lead to smaller overall effects of CPC+ than we would observe if some or all of the other initiatives did not exist. This weakening of effects would occur if the other initiatives duplicate some of the incentives and supports provided through CPC+ and these incentives and supports lead to better outcomes.

In this Appendix, we quantify how participation in other initiatives differed between CPC+ and comparison practices and how this participation shifted from the baseline period to the five program years (PYs) of CPC+ for both practice groups. To do this, we used a difference-in-differences strategy, when possible, to examine changes in participation over time between the two groups. For initiatives for which we do not have baseline period data, we examined differences between the two groups in participation during the intervention periods.

What's new this year?

1. Additional year of data (2021)
2. Results on three additional initiatives:
   - Primary Care First
   - Direct Contracting
   - Million Hearts
3. New statistics on participation in Medicare Shared Savings Program (SSP) tracks with downside risk

We were able to measure and analyze participation in five broad types of CMS initiatives through the end of CPC+: (1) Medicare fee-for-service (FFS) care management codes, (2) other Medicare FFS value-based purchasing models, (3) other primary care transformation initiatives, (4) bundled payment initiatives, and (5) population health initiatives. In Table 5.F.1, we list the specific initiatives we report results for within these five broad types, whether CPC+ practices (or their CMS-attributed Medicare FFS beneficiaries) could participate in these initiatives during the periods we study, the data source, and the definition of a beneficiary being exposed to the initiative.[95]

[95] We report whether CMS-attributed Medicare FFS beneficiaries could participate in the initiative to provide context on the level of participation expected for the CPC+ group. However, later we measure participation using the intent-to-treat evaluation sample of beneficiaries to ensure comparability between the CPC+ and comparison groups.

In addition to initiatives listed in the table, we explored participation in the following 10 initiatives: Community-Based Care Transitions Program, Comprehensive Joint Replacement model, Oncology Care Model, Independence at Home demonstration, Financial Alignment Initiative Demonstration for Medicare-Medicaid Enrollees, Comprehensive ESRD Care model, ESRD Treatment Choices Model, Value in Opioid Use Disorder Treatment model, Psychiatric Collaborative Care Management codes, and General Behavioral Health Integration. We did not include results for these initiatives because participation rates were less than 1 percent in all cases, so there was little potential either for interaction effects with CPC+ or for potentially attenuating the impacts of CPC+.

Table 5.F.1. Potential participation and our sample definition for participation in other initiatives

Type of initiative	Name of initiatives	Could active CPC+ practices or their CMS-attributed Medicare beneficiaries participate…		Data source	Definition of a beneficiary being exposed to the initiative
		During baseline period?	During intervention period?		
Medicare FFS care management codes	Chronic Care Management	Yes	No	Medicare FFS physician and outpatient claims	Beneficiary's physician billed at least one of these care management services in the year
	Transitional Care Management	Yes	Yes		
	Other care management[a]	Yes	Yes		
Other Medicare FFS value-based purchasing models	Medicare Shared Savings Program	Yes[c]	Yes[c]	CMS Master Data Management System	Beneficiary's assigned practice was in the initiative in the year,[b] or beneficiary was attributed to the initiative in the year
	Next Generation (Next Gen) ACO	No[c]	No[c]		
Other primary care transformation initiatives	Accountable Health Communities (May 2017–April 2022)	No	Yes	CMS beneficiary rosters	Beneficiary was attributed to the initiative in the year
	Transforming Clinical Practice Initiative (September 2015–September 2019)	Yes	No	CMS practitioner rosters	Beneficiary's assigned practice was in the initiative during the year[b]
	Primary Care First (January 2021–)	No	No	CMS Master Data Management System	Beneficiary's assigned practice was in the initiative in the year,[b] or beneficiary was attributed to the initiative in the year
	Direct Contracting (April 2021–)	No	No		
Bundled payment initiatives	Bundled Payment for Care Improvement (April 2013–September 2018)	Yes	Yes	Non-claims-based payment file	Beneficiary had at least one payment for a covered service in the year
	Bundled Payment for Care Improvement Advanced (October 2018–)	No	Yes		
Population health initiatives	Million Hearts (January 2017–December 2021)	No	Yes	Non-claims-based payment file[d]	Beneficiary had at least one payment for a covered service in the year

Table 5.F.1. ***(continued)***

Notes: For initiatives that started after the baseline period and/or ended before the end of CPC+, we indicated the start date and/or end date of the performance period under the name of the initiative in parentheses. In addition to initiatives listed above, we explored participation in the following initiatives: Community-Based Care Transitions Program, Comprehensive Joint Replacement model, Oncology Care Model, Independence at Home demonstration, Financial Alignment Initiative Demonstration for Medicare-Medicaid Enrollees, Comprehensive ESRD Care model, ESRD Treatment Choices Model, Value in Opioid Use Disorder Treatment model, Psychiatric Collaborative Care Management codes, and General Behavioral Health Integration. We did not include results for these initiatives because participation rates were less than 1 percent in all cases, so there was little potential either for interaction effects with CPC+ or for potentially attenuating the impacts of CPC+.

[a] This includes CPT codes G0181 (physician supervision of a Home Health Agency patient, patient not present), G0182 (physician supervision of a hospice patient, patient not present), G0502-G0504, G2214 and 99492-99494 (Psychiatric Collaborative Care Management Services), G0505 and 99483 (cognitive and function assessment for a patient with cognitive impairment), G0511 (General Care Management Services for use by RHCs or FQHCs), G0512 (Psychiatric Collaborative Care Management Services for use by RHCs or FQHCs), 99497 (advance care planning), and G2064-G2065 (Principal Care Management Services). These codes capture some type of care management but are not chronic care management or transitional care management codes.

[b] We defined a practice as being in the initiative if any of its practitioners were in the initiative.

[c] To be consistent with baseline matching, where SSP and Next Gen participation were defined as participating as of January 1, 2017, we defined baseline participation for SSP and Next Gen as participating as of January 1, 2017; for CPC+ PY 1 as participating as of January 1, 2018; for CPC+ PY 2 as participating as of January 1, 2019; for CPC+ PY 3 as participating as of January 1, 2020; and for CPC+ PY 4 as participating as of January 1, 2021. Because CPC+ ended in 2021, we defined participation for CPC+ PY 5 as participating as of December 31, 2021. CMS did not permit active CPC+ practices to participate in Next Gen as of January 1, 2017.

[d] When this report was written, the non-claims-based payment file had a complete set of payments for Million Hearts through the first four program years of CPC+ but not for the final program year.

ACO = Accountable Care Organization; CMS = Centers for Medicare & Medicaid Services; ESRD = end-stage renal disease; FFS = fee-for-service; FQHC = Federally Qualified Health Center; PY = Program Year; RHC = Rural Health Clinic; SSP = Medicare Shared Savings Program.

In the rest of this Appendix, we present the key takeaways of the results (Section 5.F.1), describe the methods used (Section 5.F.2), and discuss the results in greater detail for CPC+ practices and their matched comparison practices (Section 5.F.3). We then discuss the implications of the results for the impact analyses (Section 5.F.4).

5.F.1. Key takeaways

- In each of the five program years, both CPC+ and comparison practices had high participation in SSP—around 50 percent.
- In all other initiatives, participation was lower (each less than 13 percent). Billing for care management services was up to 24 percent among high-risk beneficiaries.
- Changes in SSP participation of CPC+ practices differed substantially from those of comparison practices. Specifically, in each of the five program years, comparison practices were more likely to participate in SSP than CPC+ practices.
  - By the end of CPC+ (PY 5), comparison practices were more likely to participate in SSP than CPC+ practices by 11.4 percentage points in Track 1 and 6.2 percentage points in Track 2 based on the practitioner-level master data management (MDM) system. The corresponding differences based on the beneficiary-level MDM were 9.2 percentage points in Track 1 and 2.8 percentage points in Track 2.
  - This was driven by CPC+ practices decreasing their participation in SSP in Track 1, and comparison practices increasing their participation in Track 2. Most of these changes happened in the years prior to PY 3, and in fact, from PY 3 to PY 5, the gap in participation decreased by 1.2 percentage points for Track 1 and by 3.2 percentage points for Track 2 based on the practitioner-level MDM. The corresponding changes based on the beneficiary-level MDM were 1.8 percentage points for Track 1 and 3.7 percentage points for Track 2.
  - These results suggest that more CPC+ practices might have chosen to participate in SSP (which is an established CMS program) if CPC+ did not exist.
  - If SSP encourages care delivery changes in the comparison group similar to those occurring in the CPC+ group, and the changes improve outcomes, we may observe only small effects of CPC+ or none at all, even if the broader model of care transformation is indeed effective in improving quality or lowering costs. As a result of these findings, we have added a sensitivity analysis to the impact modeling in which we control for contemporaneous SSP participation, to shed light on whether our impact results are at all driven by SSP participation.
  - Because some practices in CPC+ and in the comparison group started or stopped participating in SSP after CPC+ began, the findings from the impact analysis for the SSP subgroup, which is defined based on SSP status at baseline only, should be interpreted with caution. Instead of interpreting the SSP subgroup results as the impact of CPC+ combined with SSP throughout the intervention period, they should be interpreted as the impact of starting CPC+ while participating in SSP.
- Among SSP participants, both CPC+ and comparison practices increased their participation in the tracks with downside risk from 10 to 17 percent in the baseline year to 35 to 52 percent by the end of CPC+ (depending on track and CPC+ status).

- For all the other initiatives, changes in participation of CPC+ practices were similar to those of comparison practices, which suggests that differential participation in initiatives between the CPC+ and comparison groups is unlikely to influence the impact estimates.[96]

Below we describe additional key findings for CPC+ practices and their matched comparison practices over the five program years for each type of initiative.

A. Medicare FFS care management codes

- Both CPC+ and comparison practices billed any type of Medicare FFS care management codes for fewer than 13 percent of patients and had similar, small increases from baseline to the five program years of CPC+ (1 to 4 percentage points for CPC+ practices and 2 to 6 percentage points for comparison practices, depending on the program year and track).[97] Difference-in-differences estimates ranged from -1 to -2 percentage points, depending on the program year and track.
  - Both CPC+ and comparison practices billed a higher proportion of high-risk patients for care management services than for all patients, but both sets of practices still had similar changes over time. Difference-in-differences estimates ranged from -1 to -3 percentage points, depending on the program year and track.

B. Other Medicare FFS value-based purchasing models

- Comparison practices increased their participation in Medicare FFS value-based purchasing models during the intervention period by 2 to 12 percentage points (depending on initiative, track, and program year), whereas CPC+ practices either decreased or increased their participation by less than the comparison group depending on the track and the specific initiative. Difference-in-differences estimates ranged from -1 to -13 percentage points, depending on the initiative, program year, and track.
- Among SSP participants, participation in SSP tracks with downside risk grew over the five program years of CPC+ from 10 to 17 percent to 35 to 52 percent (depending on track and CPC+ status), with a large shift in growth between PY 2 and PY 3. Participation grew more among comparison practices (than among CPC+ practices) for Track 1 and grew more among CPC+ practices (than among comparison practices) for Track 2.

C. Other primary care transformation initiatives

- Medicare FFS beneficiaries in CPC+ and comparison practices had less than 2 percent participation in Accountable Health Communities (AHCs) in PY 2 (the first year of the model that beneficiaries were attributed) through PY 5.
- Reflecting CPC+ eligibility rules, CPC+ practices had much lower participation (4 to 10 percentage points lower) in the Transforming Clinical Practice Initiative (TCPI) relative to the comparison group

[96] One exception is the Transforming Clinical Practice Initiative (TCPI), where we found differential changes in participation between CPC+ and comparison practices of 2 to 8 percentage points (depending on track and program year). However, because the program ended in 2019, the differential changes were unlikely to influence the impact estimates for PY 4 and PY 5.

[97] Note that CPC+ practices were unable to bill for chronic care management codes during the model period for previously attributed patients, though they were able to bill transitional care management codes and other care management codes.

in PY 2 and PY 3 (the last performance year of TCPI).[98] Participation in TCPI decreased for both CPC+ and comparison practices from baseline to PY 2 and PY 3 but the decreases were larger for CPC+ practices than for comparison practices. Difference-in-differences estimates ranged from -2 to -8 percentage points, depending on track and program year. The performance period for TCPI ended in 2019.

- Medicare FFS beneficiaries in CPC+ and comparison practices had less than 3 percent participation in Primary Care First (PCF) and less than 4 percent in Direct Contracting (DC) in PY 5 (the first performance year of the two models).

D. Bundled payment initiatives

- Medicare FFS beneficiaries in CPC+ practices had less than 2 percent participation in Bundled Payment for Care Improvement (BPCI) at baseline, and their participation decreased further during the model period through PY 2 (the last performance year of BPCI). The comparison group had participation rates and changes similar to those of CPC+ beneficiaries.
- Medicare FFS beneficiaries in CPC+ practices had less than 1 percent participation in BPCI Advanced in PY 2 (the first performance year of the model). Participation increased by less than 2 percentage points through PY 3 and decreased afterwards through PY 5 to levels close to those in PY 2. The comparison group had participation rates and changes similar to those of CPC+ beneficiaries.

E. Population health initiatives

- Medicare FFS beneficiaries in CPC+ and comparison practices had less than 2 percent participation in Million Hearts in PY 1 (the first year of the model that beneficiaries were attributed), and their participation decreased during the model period through PY 4 (the most recent year of available data).

5.F.2. Methods

A. Measuring participation in each initiative

Overview. Although CMS provides initiatives at the practice, practitioner, and beneficiary levels, we report participation in all initiatives as the percentage of beneficiaries in each group—CPC+ and comparison—who are exposed to that initiative, separately for Track 1 and Track 2 practices. We chose to measure participation as the percentage of beneficiaries who participated because our impact estimates are at the beneficiary level. To the extent that participation in other initiatives affected the impact findings, this would likely depend on the number of beneficiaries affected by such participation. Also, reporting participation at the beneficiary level for all initiatives enables us to keep the measurements consistent across initiatives in this participation analysis.[99]

[98] Although CPC+ practices were technically unable to participate in TCPI during the CPC+ model period, we found low but non-zero participation rates among CPC+ practices (2.6 and 2.7 percent for Track 1 in PYs 2 and 3, respectively, and 2 percent for Track 2 in both PYs 2 and 3), which may be explained by belated withdrawals, differences between the IQVIA and CMS practitioner rosters, or the intent-to-treat approach, which continues to follow practices that no longer participate in CPC+.

[99] For some initiatives, like CCM, participation is inherently at the beneficiary level since billing for CCM services occurs on a per-beneficiary basis. However, for other initiatives, like TCPI, Next Gen, and SSP, practices decide whether or not to participate, and we assume that all beneficiaries assigned to participating practices were affected. Also, we selected comparison practices based on baseline participation in SSP weighted at the beneficiary level. Therefore, we assess the balance in CPC+ and comparison practices' SSP participation at that level.

Beneficiary-level initiatives. We measured provision of Medicare FFS care management services as the percentage of beneficiaries whose practitioner billed for at least one of those services in that year. We also looked at participation in Medicare FFS care management services for high-risk beneficiaries, defined as beneficiaries who had a hierarchical condition category (HCC) score greater than the 90th percentile of the distribution of HCC scores among assigned beneficiaries within their track or had Alzheimer's disease or dementia (indicated by the Chronic Conditions Warehouse) in 2015 for baseline and 2016 for intervention periods, because care management services are targeted to high-risk beneficiaries. We measured participation in AHC as the percentage of beneficiaries who were attributed to organizations participating in AHC based on CMS beneficiary rosters. We measured participation in BPCI, BPCI Advanced, and Million Hearts as the percentage of beneficiaries who had at least one payment for a covered service in the year based on non-claims-based payment files from CMS.

Practitioner-level initiatives. Since models such as SSP, Next Gen, TCPI, PCF, and DC report practitioners' participation in the initiatives, as opposed to practice sites participating, we first used the IQVIA practitioner roster to roll practitioner participation up to the practice site level by counting a practice as participating if any practitioner in the practice was reported as participating.[100] We then weighted practice participation by the number of Medicare beneficiaries assigned to that practice in the baseline year so we can interpret the results as the number of beneficiaries who were participating in the initiative.[101] Inferring beneficiary participation from practitioner participation tends to inflate participation because a practice and all of its assigned beneficiaries are determined to be participating in the model as long as the practice had at least one participating practitioner.[102] As a robustness check, we also used the beneficiary-level MDM to directly measure beneficiary participation (rather than inferring beneficiary participation from practitioner-level participation) in SSP, Next Gen, PCF, and DC. Among SSP participants, we flagged each practice as participating in downside risk if at least one of their practitioners participated in an SSP track that involves sharing the risk of losses.[103]

Timing of measured participation. For initiatives identified from the practitioner-level or beneficiary-level MDM (that is, SSP, Next Gen, PCF, and DC), we measured participation for each program year as of January 1 of the following calendar year, which is consistent with how we defined SSP participation at baseline for the main impact evaluation (which was as of January 1, 2017). For example, PY 1 participation was defined as of January 1, 2018. Similarly, PY 4 participation was defined as of January 1, 2021. Because CPC+ ended in 2021, we defined participation for CPC+ PY 5 as participating as of December 31, 2021. For all other initiatives, we measured participation in the respective program year.

[100] The MDM reports 90 percent of participation in SSP at the Tax Identification Number (TIN) level, and 10 percent at the National Provider Identifier (NPI)/TIN level. Since TINs are not unique at the practice level, we merged measures of participation of all practitioners to whom we assigned that TIN, and then rolled up participation to the practice level using the IQVIA practitioner roster.

[101] This is the same method that we used for comparison selection. That is, we first looked at practitioner-level participation in SSP or other initiatives and then rolled these measures up to the practice level. Then, we weighted by the number of beneficiaries in the practice in the baseline year.

[102] That is, practices in which some or all of the practitioners participated in one of these models would equally be considered as participating in the model.

[103] The SSP tracks that involve downside risk include Track 1+, Track 2, Track 3, Basic C, Basic D, Basic E, and Enhanced; the SSP tracks that do not involve downside risk include Track 1, Basic A, and Basic B.

B. Analytic approach

Overview. To estimate difference-in-differences changes in participation in each initiative, comparing the CPC+ and comparison practices from the baseline year through the end of CPC+, we used a regression model similar to the one used for all claims-based beneficiary-level outcomes described in this report (see Chapter 5 in O'Malley et al. 2023), but we did not include any additional regression covariates other than the difference-in-differences estimators. We did not include additional controls since the goal of the analysis was to understand the total, non-adjusted participation in initiatives.

Intent-to-treat (ITT) approach. As in the impact evaluation, we conducted analyses using the ITT sample of beneficiaries, to ensure comparability between the CPC+ and comparison groups over the evaluation period. Under the ITT approach, beneficiaries are assigned to the first CPC+ or comparison practice to which they were attributed in the baseline or follow-up period, even if they began seeing a different primary care practice more frequently later in that period (as long as they satisfy the eligibility criteria). In addition, we follow all CPC+ practices that were participating in CPC+ as of April 2017. As a result, the sample includes practices that no longer participate in CPC+ or beneficiaries who are no longer attributed to a CPC+ practice.

Level of regressions. We conducted analyses at two levels—beneficiary-level and practice-level. In both analyses, we can interpret participation as the percentage of beneficiaries who were participating in the initiative.

- *Beneficiary-level analyses.* For the initiatives that had observations at the beneficiary level, we ran regressions at the beneficiary level, and used beneficiary-level matching weights.

- *Practice-level analyses.* For the initiatives for which we rolled up participation to the practice level, we estimated regressions at the practice level and used practice-level matching weights that also weight practices by the number of beneficiaries in that practice during the baseline period. Therefore, the results can be interpreted as the number of beneficiaries who were participating in the initiative.

Initiatives with incomplete data. For AHC, BPCI Advanced, PCF, DC, and Million Hearts, we present the participation rates and the percentage point differences in each program year, but not the difference-in-differences changes because the initiatives were not present at baseline. For BPCI, we present data through PY 2 because it ended in 2018; for Million Hearts, we present data through PY 4, the most recent year of available data; for TCPI, we present data through PY 3 because it ended in 2019.

5.F.3. Results over the five program years

Table 5.F.2 reports participation of beneficiaries in various initiatives by time period (baseline year and PY 1 through PY 5) for CPC+ practices and their comparison practices for Track 1 and Track 2, respectively. In this table, years in which we did not have data or in which the initiative was not active are not shown. For example, initiatives that began in 2017 or after have no rows for the year before the initiative started. Similarly, initiatives that ended prior to PY 5 have no rows from the year after the initiatives ended through PY 5.

Table 5.F.2. Participation in other initiatives by beneficiaries in CPC+ practices and comparison practices in the baseline and five program years, Track 1 and Track 2

Total beneficiary participation in CMS initiatives was high for SSP, but less than 13 percent for all other initiatives. Comparison practices had participation similar to that of CPC+ practices over time, except for SSP and TCPI, for which participation grew by at least 5 percentage points more among comparison practices than among CPC+ practices in certain program years.

		Track 1				Track 2			
	Time period	Percentage of Medicare FFS beneficiaries exposed to the initiative, CPC+ group	Percentage of Medicare FFS beneficiaries exposed to the initiative, Comparison group	Percentage point difference	Percentage point difference-in-differences estimate (90% CI)	Percentage of Medicare FFS beneficiaries exposed to the initiative, CPC+ group	Percentage of Medicare FFS beneficiaries exposed to the initiative, Comparison group	Percentage point difference	Percentage point difference-in-differences estimate (90% CI)
Type of initiative: Medicare FFS care management codes									
Name of initiative									
Chronic Care Management									
All beneficiaries	Baseline	1.1	1.6	-0.5	n.a.	1.5	1.9	-0.5	n.a.
	PY 1	0.7	2.7	-2.0	-1.5*** (-1.8, -1.2)	0.7	2.5	-1.8	-1.3*** (-1.7, -1.0)
	PY 2	1.1	2.9	-1.8	-1.4*** (-1.7, -1.1)	1.2	3.0	-1.8	-1.3*** (-1.7, -0.9)
	PY 3	1.3	3.3	-2.0	-1.5*** (-1.8, -1.2)	1.4	3.4	-2.0	-1.6*** (-2.0, -1.1)
	PY 4	1.8	4.1	-2.3	-1.8*** (-2.2, -1.4)	1.7	4.2	-2.5	-2.0*** (-2.6, -1.5)
	PY 5	1.8	4.0	-2.1	-1.6*** (-2.1, -1.2)	1.7	4.1	-2.3	-1.8*** (-2.4, -1.3)
High-risk beneficiaries[a]	Baseline	2.2	2.9	-0.8	n.a.	3.0	4.2	-1.2	n.a.
	PY 1	1.6	4.9	-3.2	-2.5*** (-3.0, -2.0)	1.7	5.2	-3.5	-2.3*** (-3.0, -1.6)
	PY 2	2.5	5.8	-3.2	-2.5*** (-3.0, -2.0)	2.8	6.3	-3.5	-2.3*** (-3.1, -1.5)
	PY 3	3.2	6.8	-3.6	-2.8*** (-3.3, -2.3)	3.2	7.3	-4.1	-2.9*** (-3.9, -1.9)
	PY 4	4.4	8.5	-4.0	-3.3*** (-4.0, -2.6)	4.2	8.8	-4.6	-3.4*** (-4.8, -2.1)
	PY 5	4.9	8.8	-3.9	-3.1*** (-4.0, -2.3)	4.7	8.8	-4.2	-2.9*** (-4.4, -1.4)
Transitional Care Management									
All beneficiaries	Baseline	3.7	3.4	0.3	n.a.	4.8	3.4	1.3	n.a.
	PY 1	4.6	3.8	0.8	0.5*** (0.4, 0.7)	5.3	3.8	1.5	0.1 (0.0, 0.3)
	PY 2	5.4	4.2	1.2	0.9*** (0.6, 1.1)	5.8	4.2	1.6	0.2** (0.1, 0.4)
	PY 3	5.7	4.7	1.1	0.8*** (0.5, 1.0)	6.1	4.7	1.4	0.1 (-0.1, 0.3)
	PY 4	4.4	3.9	0.5	0.2 (0.0, 0.4)	4.8	4.0	0.8	-0.6*** (-0.8, -0.4)
	PY 5	4.6	4.3	0.4	0.1 (-0.2, 0.3)	5.0	4.4	0.7	-0.7*** (-0.9, -0.5)
High-risk beneficiaries[a]	Baseline	8.7	7.6	1.0	n.a.	11.0	8.0	3.0	n.a.
	PY 1	10.6	8.8	1.8	0.7*** (0.3, 1.1)	12.3	8.8	3.4	0.4* (0.1, 0.8)
	PY 2	11.4	9.1	2.3	1.3*** (0.7, 1.8)	12.6	9.2	3.4	0.4 (-0.1, 0.9)
	PY 3	12.1	9.9	2.3	1.2*** (0.6, 1.8)	12.9	10.0	3.0	-0.1 (-0.6, 0.5)
	PY 4	9.4	8.3	1.1	0.0 (-0.6, 0.6)	10.2	8.5	1.7	-1.3*** (-1.9, -0.7)
	PY 5	10.0	9.0	1.0	0.0 (-0.6, 0.6)	10.7	9.1	1.6	-1.5*** (-2.0, -0.9)

Table 5.F.2. ***(continued)***

		Track 1				Track 2			
	Time period	Percentage of Medicare FFS beneficiaries exposed to the initiative, CPC+ group	Percentage of Medicare FFS beneficiaries exposed to the initiative, Comparison group	Percentage point difference	Percentage point difference-in-differences estimate (90% CI)	Percentage of Medicare FFS beneficiaries exposed to the initiative, CPC+ group	Percentage of Medicare FFS beneficiaries exposed to the initiative, Comparison group	Percentage point difference	Percentage point difference-in-differences estimate (90% CI)
Other care management[b]									
All beneficiaries	Baseline	2.9	2.0	0.9	n.a.	2.7	2.2	0.5	n.a.
	PY 1	3.7	3.2	0.4	-0.5* (-0.9, 0.0)	3.8	3.3	0.5	0.1 (-0.3, 0.4)
	PY 2	4.1	4.1	0.0	-0.8*** (-1.4, -0.3)	4.6	4.2	0.4	-0.1 (-0.7, 0.5)
	PY 3	4.8	5.1	-0.3	-1.2*** (-1.7, -0.6)	5.6	4.9	0.7	0.2 (-0.5, 0.9)
	PY 4	5.1	5.4	-0.2	-1.1*** (-1.7, -0.6)	5.5	5.0	0.5	0.1 (-0.6, 0.7)
	PY 5	5.9	5.9	0.0	-0.9** (-1.5, -0.3)	6.2	5.6	0.6	0.1 (-0.6, 0.8)
High-risk beneficiaries[a]	Baseline	4.4	3.8	0.6	n.a.	4.1	4.3	-0.2	n.a.
	PY 1	6.0	6.0	0.0	-0.6** (-1.1, -0.1)	6.0	6.0	0.0	0.2 (-0.4, 0.7)
	PY 2	7.2	7.5	-0.3	-0.9** (-1.5, -0.3)	7.6	7.5	0.2	0.4 (-0.4, 1.1)
	PY 3	8.9	9.5	-0.5	-1.1*** (-1.8, -0.5)	9.9	9.2	0.7	0.9* (0.1, 1.7)
	PY 4	9.5	10.1	-0.5	-1.1*** (-1.8, -0.5)	10.1	9.8	0.3	0.5 (-0.3, 1.3)
	PY 5	10.9	11.5	-0.6	-1.1** (-2.0, -0.3)	11.3	11	0.3	0.5 (-0.4, 1.5)
Combined measure of care management services									
Any care management[c]									
All beneficiaries	Baseline	7.2	6.4	0.8	n.a.	8.4	6.9	1.5	n.a.
	PY 1	8.5	8.7	-0.2	-1.0*** (-1.5, -0.5)	9.3	8.7	0.6	-0.9*** (-1.4, -0.4)
	PY 2	9.8	9.9	-0.1	-0.9** (-1.5, -0.3)	10.8	10.1	0.6	-0.9** (-1.5, -0.2)
	PY 3	10.9	11.4	-0.6	-1.3*** (-2.0, -0.7)	11.9	11.4	0.5	-1.1** (-1.8, -0.3)
	PY 4	10.3	11.6	-1.3	-2.1*** (-2.8, -1.4)	10.8	11.4	-0.6	-2.1*** (-2.9, -1.3)
	PY 5	11.1	12.1	-1.0	-1.8*** (-2.5, -1.1)	11.6	12.0	-0.4	-1.9*** (-2.8, -1.1)
High-risk beneficiaries[a]	Baseline	14.0	13.1	1.0	n.a.	16.5	14.4	2.1	n.a.
	PY 1	16.7	17.2	-0.5	-1.5*** (-2.1, -0.8)	18.2	17.5	0.7	-1.4*** (-2.2, -0.6)
	PY 2	19.0	19.2	-0.1	-1.1** (-2.0, -0.3)	20.6	19.5	1.1	-1.1* (-2.0, -0.2)
	PY 3	21.3	21.7	-0.4	-1.3** (-2.2, -0.5)	22.7	22.1	0.6	-1.5** (-2.7, -0.4)
	PY 4	20.1	22.0	-1.9	-2.8*** (-3.8, -1.9)	21.0	22.2	-1.2	-3.3*** (-4.8, -1.8)
	PY 5	21.8	23.5	-1.7	-2.6*** (-3.7, -1.6)	22.6	23.3	-0.8	-2.9*** (-4.5, -1.2)
Type of initiative: Other Medicare FFS value-based purchasing models									
Name of initiative									
Medicare Shared Savings Program									
Practitioner-level MDM[d,e]	Baseline	51.4	52.3	-0.9	n.a.	44.2	44.2	0.0	n.a.
	PY 1	53.2	58.7	-5.5	-4.6*** (-7.5, -1.7)	44.8	53.6	-8.7	-8.7*** (-11.8, -5.7)
	PY 2	48.7	55.8	-7.1	-6.1*** (-9.7, -2.6)	41.6	51.7	-10.1	-10.1*** (-13.8, -6.4)
	PY 3	45.1	58.7	-13.6	-12.7*** (-16.6, -8.8)	46.4	55.8	-9.4	-9.4*** (-13.7, -5.1)
	PY 4	45.2	57.5	-12.3	-11.4*** (-15.4, -7.3)	48.1	54.4	-6.3	-6.3** (-10.5, -2.1)
	PY 5	44.7	57.0	-12.4	-11.4*** (-15.5, -7.4)	47.7	53.9	-6.2	-6.2** (-10.5, -1.9)

Table 5.F.2. ***(continued)***

		Track 1				Track 2			
	Time period	Percentage of Medicare FFS beneficiaries exposed to the initiative, CPC+ group	Percentage of Medicare FFS beneficiaries exposed to the initiative, Comparison group	Percentage point difference	Percentage point difference-in-differences estimate (90% CI)	Percentage of Medicare FFS beneficiaries exposed to the initiative, CPC+ group	Percentage of Medicare FFS beneficiaries exposed to the initiative, Comparison group	Percentage point difference	Percentage point difference-in-differences estimate (90% CI)
Beneficiary-level MDM[d,f]	Baseline	48.8	44.2	4.7	n.a.	41.2	38.1	3.1	n.a.
	PY 1	51.5	50.1	1.4	-3.2** (-5.7, -0.8)	42.9	46.5	-3.6	-6.7*** (-9.4, -4.1)
	PY 2	46.1	46.5	-0.4	-5.0*** (-7.9, -2.2)	39.6	43.4	-3.7	-6.9*** (-10.0, -3.7)
	PY 3	44.5	50.9	-6.4	-11.1*** (-14.2, -7.9)	44.5	47.8	-3.3	-6.5*** (-9.8, -3.1)
	PY 4	43.3	48.5	-5.2	-9.8*** (-13.0, -6.6)	45.3	45.9	-0.6	-3.8* (-7.1, -0.4)
	PY 5	42.8	47.4	-4.6	-9.2*** (-12.5, -6.0)	45.6	45.2	0.4	-2.8 (-6.1, 0.6)
Next Generation									
Practitioner-level MDM[d,g]	Baseline	0.0	0.0	0.0	n.a.	0.2	0.0	0.2	n.a.
	PY 1	0.2	3.2	-3.0	-3.0*** (-3.7, -2.2)	1.1	3.0	-2.0	-2.1*** (-3.2, -1.0)
	PY 2	0.5	4.4	-3.9	-3.9*** (-5.1, -2.6)	1.4	3.7	-2.3	-2.5*** (-3.8, -1.3)
	PY 3	0.2	3.9	-3.7	-3.7*** (-5.0, -2.5)	1.2	3.1	-1.9	-2.1*** (-3.3, -0.9)
	PY 4	1.2	3.3	-2.2	-2.2*** (-3.4, -0.9)	1.2	2.5	-1.3	-1.5** (-2.6, -0.3)
	PY 5	1.2	3.3	-2.1	-2.1*** (-3.4, -0.9)	1.2	2.3	-1.1	-1.3* (-2.4, -0.2)
Beneficiary-level MDM[d,f]									
	Baseline	0.0	0.0	0.0	n.a.	0.0	0.0	0.0	n.a.
	PY 1	0.3	3.0	-2.8	-2.8*** (-3.4, -2.2)	1.1	3.0	-1.9	-1.9*** (-2.8, -0.9)
	PY 2	0.4	3.9	-3.6	-3.6*** (-4.4, -2.7)	1.2	3.5	-2.3	-2.3*** (-3.3, -1.4)
	PY 3	0.4	3.6	-3.2	-3.2*** (-4.0, -2.3)	1.1	3.0	-1.9	-1.9*** (-2.7, -1.0)
	PY 4	1.1	3.1	-2.1	-2.1*** (-3.0, -1.2)	1.2	2.6	-1.4	-1.4*** (-2.2, -0.6)
	PY 5	1.0	3.2	-2.1	-2.1*** (-3.0, -1.2)	1.1	2.6	-1.5	-1.5*** (-2.2, -0.7)
Type of initiative: Other primary care transformation initiatives									
Name of initiative									
Accountable Health Communities	PY 2	0.1	0.3	-0.2	n.a.	0.1	0.3	-0.3	n.a.
	PY 3	0.6	0.8	-0.2	n.a.	0.6	0.8	-0.2	n.a.
	PY 4	1.1	1.0	0.1	n.a.	1.4	1.1	0.3	n.a.
	PY 5	1.3	1.2	0.2	n.a.	1.8	1.3	0.5	n.a.
Transforming Clinical Practice Initiative[h]	Baseline	10.9	10.8	0.1	n.a.	9.9	12.8	-2.9	n.a.
	PY 1	10.3	12.2	-1.8	-2.0** (-3.6, -0.3)	9.9	14.5	-4.6	-1.7** (-3.0, -0.4)
	PY 2	2.6	10.5	-7.9	-8.0*** (-10.5, -5.5)	2.0	12.1	-10.1	-7.3*** (-9.3, -5.2)
	PY 3	2.7	7.0	-4.4	-4.5*** (-7.3, -1.6)	2.0	7.4	-5.4	-2.5 (-5.2, 0.2)
Primary Care First (Practitioner-level MDM[d,g,i])	PY 5	1.8	1.4	0.4	n.a.	2.6	1.3	1.3	n.a.
Primary Care First (Beneficiary-level MDM[d,f,i])	PY 5	1.5	1.2	0.3	n.a.	1.2	1.0	0.3	n.a.

Table 5.F.2. ***(continued)***

		Track 1				Track 2			
	Time period	Percentage of Medicare FFS beneficiaries exposed to the initiative, CPC+ group	Percentage of Medicare FFS beneficiaries exposed to the initiative, Comparison group	Percentage point difference	Percentage point difference-in-differences estimate (90% CI)	Percentage of Medicare FFS beneficiaries exposed to the initiative, CPC+ group	Percentage of Medicare FFS beneficiaries exposed to the initiative, Comparison group	Percentage point difference	Percentage point difference-in-differences estimate (90% CI)
Direct Contracting (Practitioner-level MDM[d,g,i])	PY 5	1.1	1.3	-0.2	n.a.	3.9	1.3	2.6	n.a.
Direct Contracting (Beneficiary-level MDM[d,f,i])	PY 5	1.0	1.2	-0.1	n.a.	2.9	1.0	1.9	n.a.
Type of initiative: Bundled payment initiatives									
Name of initiative									
Bundled Payment for Care Improvement[j]	Baseline	1.6	1.7	-0.1	n.a.	1.7	1.8	-0.1	n.a.
	PY 1	1.3	1.4	-0.1	-0.03 (-0.1, 0.04)	1.5	1.5	0.0	0.04 (-0.02, 0.1)
	PY 2	0.9	0.9	-0.1	0.02 (-0.1, 0.1)	1.0	1.0	0.0	0.1* (0.01, 0.2)
Bundled Payment for Care Improvement – Advanced[k]	PY 2	0.5	0.4	0.1	n.a.	0.5	0.4	0.1	n.a.
	PY 3	2.1	1.6	0.5	n.a.	2.2	1.7	0.6	n.a.
	PY 4	1.4	1.3	0.1	n.a.	1.5	1.3	0.2	n.a.
	PY 5	0.5	0.5	0.1	n.a.	0.6	0.5	0.1	n.a.
Type of initiative: Population health initiatives									
Name of initiative									
Million Hearts	PY 1	0.8	0.7	0.1	n.a.	1.7	0.5	1.2	n.a.
	PY 2	0.3	0.1	0.1	n.a.	0.3	0.1	0.2	n.a.
	PY 3	0.0	0.0	0.0	n.a.	0.0	0.0	0.0	n.a.
	PY 4	0.0	0.0	0.0	n.a.	0.0	0.0	0.0	n.a.

Source: Analysis of Medicare FFS claims for 2016 through 2021; practitioner-level MDM extracts from February 27, 2017, February 23, 2018, February 26, 2019, February 28, 2020, February 26, 2021, and February 25, 2022; beneficiary-level MDM extracts from January 27, 2017, February 23, 2018, February 26, 2019, February 28, 2020, February 25, 2021, and February 25, 2022; CMS January 2017, 2018, 2019, and 2020 TCPI rosters; CMS 2022 AHC roster, and the non-claims-based payment extract, which had payments up to February 1, 2022.

Notes: We report participation in initiatives as the percentage of beneficiaries who were exposed to the initiative in each period in each track in each group (CPC+ or comparison practices), with comparison practices weighted using matching weights. Initiatives that are not at the beneficiary level are weighted by the number of beneficiaries assigned to that practice during the baseline period, so that the results can also be interpreted as the percentage of beneficiaries who were participating in the initiative. We calculated the difference in participation in a given year in each track between CPC+ and comparison practices as the percentage point difference. We calculated the difference-in-differences estimate as the difference in percentage participation between CPC+ and comparison practices in the relevant program period (PY 1 through PY 5), minus the difference in the baseline period. The difference-in-differences estimate is in percentage point units. We estimated 90 percent confidence intervals calculating standard errors using linear regression and clustering at the practice level. n.a. indicates that the difference-in-differences estimate is not applicable, because we do not have data for the baseline period. 0.0 indicates that <0.05 percent of beneficiaries participated in the initiative, or that there was a less than 0.05 percentage point difference in participation between CPC+ and comparison practices. Note that the percentage point difference and the percentage point difference-in-differences estimates shown may differ from the corresponding calculations based on the percentages in the cells due to rounding. For Medicare FFS care management codes, the population we used to calculate participation is indicated under the name of the initiative in parentheses. For the rest of the initiatives, we used the full population.

Table 5.F.2. ***(continued)***

*/**/*** Significantly different from zero at the 0.10/0.05/0.01 level, two-tailed test.

[a] We defined high-risk beneficiaries as those who had an HCC score greater than the 90th percentile of the distribution of HCC scores among assigned beneficiaries within their track or had Alzheimer's disease or dementia as indicated by the Chronic Conditions Warehouse. For baseline, we calculated HCC scores from 2015 claims. For the intervention period, we calculated HCC scores from 2016 claims.

[b] This includes CPT codes G0181 (physician supervision of a Home Health Agency patient, patient not present), G0182 (physician supervision of a hospice patient, patient not present), G0502-G0504, G2214 and 99492-99494 (Psychiatric Collaborative Care Management Services), G0505 and 99483 (cognitive and function assessment for a patient with cognitive impairment), G0511 (General Care Management Services for use by RHCs or FQHCs), G0512 (Psychiatric Collaborative Care Management Services for use by RHCs or FQHCs), 99497 (advance care planning), and G2064-G2065 (Principal Care Management Services). These codes capture some type of care management but are not chronic care management or transitional care management codes.

[c] This includes beneficiaries whose physicians billed at least one chronic care management, transitional care management, or other care management service.

[d] The date used to define whether a practice participated in SSP, Next Gen, PCF, and DC at baseline was January 1, 2017 (consistent with the date used to define participation in comparison group selection). Accordingly, we defined the PY 1 participation value as participation as of January 1, 2018, the PY 2 participation value as participation as of January 1, 2019, the PY 3 participation value as participation as of January 1, 2020, and the PY 4 participation value as participation as of January 1, 2021. Because CPC+ ended in 2021, we defined participation for CPC+ PY 5 as participating as of December 31, 2021.

[e] In the practitioner MDM, 91 percent of participation in SSP is counted at the TIN level, while the remaining 9 percent is at the NPI-TIN level. If an NPI was listed in the practitioner MDM, we counted all practices with an NPI-TIN listed in that year as participating in SSP. If the NPI was missing in the practitioner MDM, we counted all practices with the TIN listed in that year as participating in SSP.

[f] In the beneficiary MDM, participation is at the beneficiary level, and we measured participation as the fraction of beneficiaries in each sample (i.e., CPC+ and comparison group practices) who participated in the initiative. Because inferring beneficiary participation from practitioner participation tends to inflate participation, we separately measured participation based on the beneficiary MDM as a robustness check for the measure of participation based on the practitioner MDM.

[g] In the practitioner MDM, participation in Next Gen, PCF, and DC is at the NPI-TIN level. We counted all practices with an NPI-TIN listed in that year as participating in Next Gen, PCF, or DC.

[h] CPC+ practices were technically unable to participate in TCPI during the CPC+ intervention period; however, we found that 10.3 percent of CPC+ practices did not withdraw from TCPI before the beginning of 2017. This is likely because the practices did not immediately initiate withdrawal. For PY 2 and PY 3, we also found lower but non-zero participation rates among CPC+ practices (2.6 and 2.7 percent), which may be explained by additional belated withdrawals, differences between the IQVIA and CMS practitioner rosters, or the intent-to-treat approach, which continues to follow practices that no longer participate in CPC+. We do not have participation data starting in 2020 (i.e., PY 4) because TCPI ended in September 2019.

[i] PCF and DC began in January 2021 and April 2021, respectively, both in PY 5. CPC+ practices were technically unable to participate in PCF or DC during the CPC+ intervention period; however, we observed non-zero participation in PCF and DC among CPC+ practices, which may be explained by differences between the IQVIA and CMS practitioner rosters, or the intent-to-treat approach, which continues to follow practices that no longer participate in CPC+ or beneficiaries who are no longer attributed to a CPC+ practice.

[j] We measured participation based on the non-claims-based payment extract. We do not have participation data starting in 2019 (i.e., PY 3) because BPCI ended in September 2018.

[k] BPCI Advanced began in October 2018 (i.e., PY 2). We measured participation based on the non-claims-based payment extract.

ACO = Accountable Care Organization; AHC = Accountable Health Communities; BPCI = Bundled Payment for Care Improvement; CI = confidence interval; CMS = Centers for Medicare & Medicaid Services; CPT = Current Procedural Terminology; DC = Direct Contracting; FFS = fee-for-service; FQHC = Federally Qualified Health Center; HCC = hierarchical condition category; MDM = CMS Master Data Management System; n.a. = not applicable; NPI = National Provider Identifier; PCF = Primary Care First; PY = Program Year; RHC = Rural Health Clinic; SSP = Medicare Shared Savings Program; TCPI = Transforming Clinical Practice Initiative; TIN = Taxpayer Identification Number.

A. Billing for Medicare FFS care management services

Less than 13 percent of all beneficiaries had claims for Medicare FFS care management services and relative changes from the baseline period to the five years of CPC+ between CPC+ and comparison practices were less than 3 percentage points.

- Between 7 and 12 percent of CPC+-assigned[104] Medicare FFS beneficiaries and between 6 and 13 percent of comparison beneficiaries had claims for at least one of the care management service types (transitional care management [TCM], CCM, or other care management) over the six years we examined.
- Less than 7 percent of Medicare FFS beneficiaries had claims for each particular type of these services over the six years we examined.
- CPC+ and comparison practices experienced small changes over time.
  - From the baseline to the five years of CPC+, CPC+ practices had a less than 1 percentage point change in their billing for CCM services, and comparison practices increased their billing for CCM services by less than 3 percentage points.
  - CPC+ practices increased their billing for TCM services by 0.1 to 0.9 percentage points more than comparison practices in the first three program years. The trend shifted starting in PY3.
    - From PY 3 to PY 4, both CPC+ and comparison practices decreased TCM billing by 0.7 to 1.3 percentage points, but the decrease was slightly larger for CPC+ practices than for comparison practices, resulting in billing rates becoming more similar between CPC+ and comparison practices. This finding may be driven by service use changes partly because of coronavirus disease 2019 (COVID-19); in our analyses of COVID-19, we found that both CPC+ and comparison regions experienced a reduction in service utilization in 2020, but the reduction was up to 2 percent larger in CPC+ regions than in comparison regions (see Appendix 5D).[105]
    - While TCM billing slightly increased from PY 4 to PY 5 for both CPC+ and comparison practices (by 0.2 to 0.4 percentage points), billing rates were still lower than those in PY 3 for both practice groups.
  - In Track 1 practices, CPC+ practices increased their billing for other care management services[106] from baseline to PY 5 by 0.9 percentage points less than comparison practices, and in Track 2, CPC+ practices increased their billing by 0.1 percentage point more than comparison practices.
  - Figure 5.F.1 shows trends in CPC+ and comparison group participation in any care management services from baseline through the end of CPC+ for Track 1 and Track 2 practices, for all

[104] Assigned Medicare FFS beneficiaries are those who are in our intent-to-treat (ITT) sample. Under our ITT approach, beneficiaries are assigned to the first CPC+ practice or comparison practice to which they were attributed in the baseline or follow-up period, even if they began seeing a different primary care practice more frequently later in that period (as long as they satisfy the eligibility criteria).

[105] Unlike the impact evaluation, we did not include any COVID-19-related controls.

[106] This includes CPT codes G0181 (physician supervision of a Home Health Agency patient, patient not present), G0182 (physician supervision of a hospice patient, patient not present), G0502-G0504, G2214, and 99492-99494 (Psychiatric Collaborative Care Management Services), G0505 and 99483 (cognitive and function assessment for a patient with cognitive impairment), G0511 (General Care Management Services for use by RHCs or FQHCs), G0512 (Psychiatric Collaborative Care Management Services for use by RHCs or FQHCs), 99497 (advance care planning), and G2064-G2065 (Principal Care Management Services). These codes capture some type of care management but are not chronic care management or transitional care management codes.

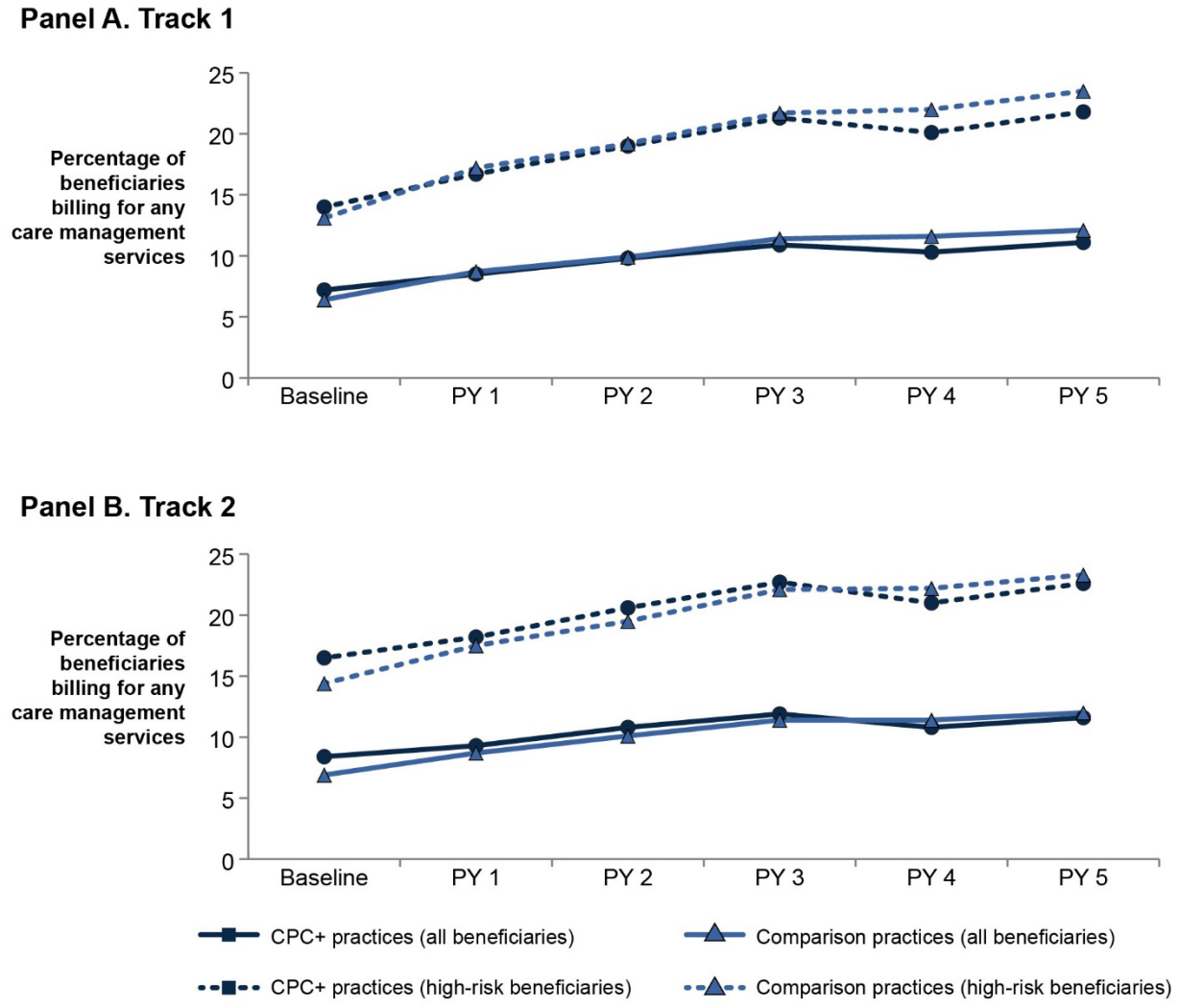
beneficiaries and for high-risk beneficiaries. The proportion of all beneficiaries who had any claims for care management services grew by 1 to 4 percentage points for CPC+ practices and 2 to 6 percentage points for comparison practices from baseline to the five program years of CPC+, with comparison practices' participation growing by 1 to 2 percentage points more than CPC+ practices for Tracks 1 and 2.

The difference-in-differences estimates are quantitatively small (less than 3 percentage points) due to low overall use of these types of claims throughout the observation period. We checked whether the low use could reflect that only a limited population of beneficiaries were eligible. However, even among high-risk beneficiaries, less than 24 percent of such beneficiaries received care management services and the difference-in-differences estimates remained less than 4 percentage points. These small differences will be unlikely to translate into substantial differences in Medicare expenditures, and thus unlikely to affect estimated impacts of CPC+.

Our analysis above focuses on differences in billable care management services between CPC+ and comparison practices, which may or may not translate into differences in the provision of total care management services (billable and non-billable). This might be particularly true for CCM services because CPC+ practices are unable to bill previously attributed beneficiaries for these services. However, findings from the CPC+ Care Delivery Reporting data indicate there has been little change in the number of CPC+ beneficiaries receiving longitudinal care management services over the course of CPC+ (see Chapter 4.3.2 of Swankoski et al. [2022] and O'Malley et al. 2023), suggesting that our findings on billable care management services might also translate into provision of total care management services.

Figure 5.F.1. Trends of billing for any care management services by beneficiaries in CPC+ practices and comparison practices in the baseline and five program years: Track 1 and Track 2

Billing for any care management was less than 13 percent for all beneficiaries and up to 24 percent for high-risk beneficiaries. Participation grew less over time among CPC+ practices than among comparison practices.



Source: Analysis of Medicare FFS claims for 2016 through 2021.

Notes: We report participation in initiatives as the percentage of beneficiaries whose physicians billed at least one chronic care management, transitional care management, or other care management service in each period in each track in each group (CPC+ or comparison practices), with comparison practices weighted using matching weights. The population we used to calculate participation is indicated in the chart legend in parentheses. We defined high-risk beneficiaries as those who had an HCC score greater than the 90th percentile of the distribution of HCC scores among assigned beneficiaries within their track or had Alzheimer's disease or dementia as indicated by the Chronic Conditions Warehouse. For baseline, we calculated HCC scores from 2015 claims. For the intervention period, we calculated HCC scores from 2016 claims.

FFS = fee-for-service; HCC = hierarchical condition category; PY = Program Year.

B. Participation in other Medicare FFS value-based purchasing models

In the five program years of CPC+, participation in other Medicare FFS value-based purchasing models grew among comparison practices relative to CPC+ practices.[107]

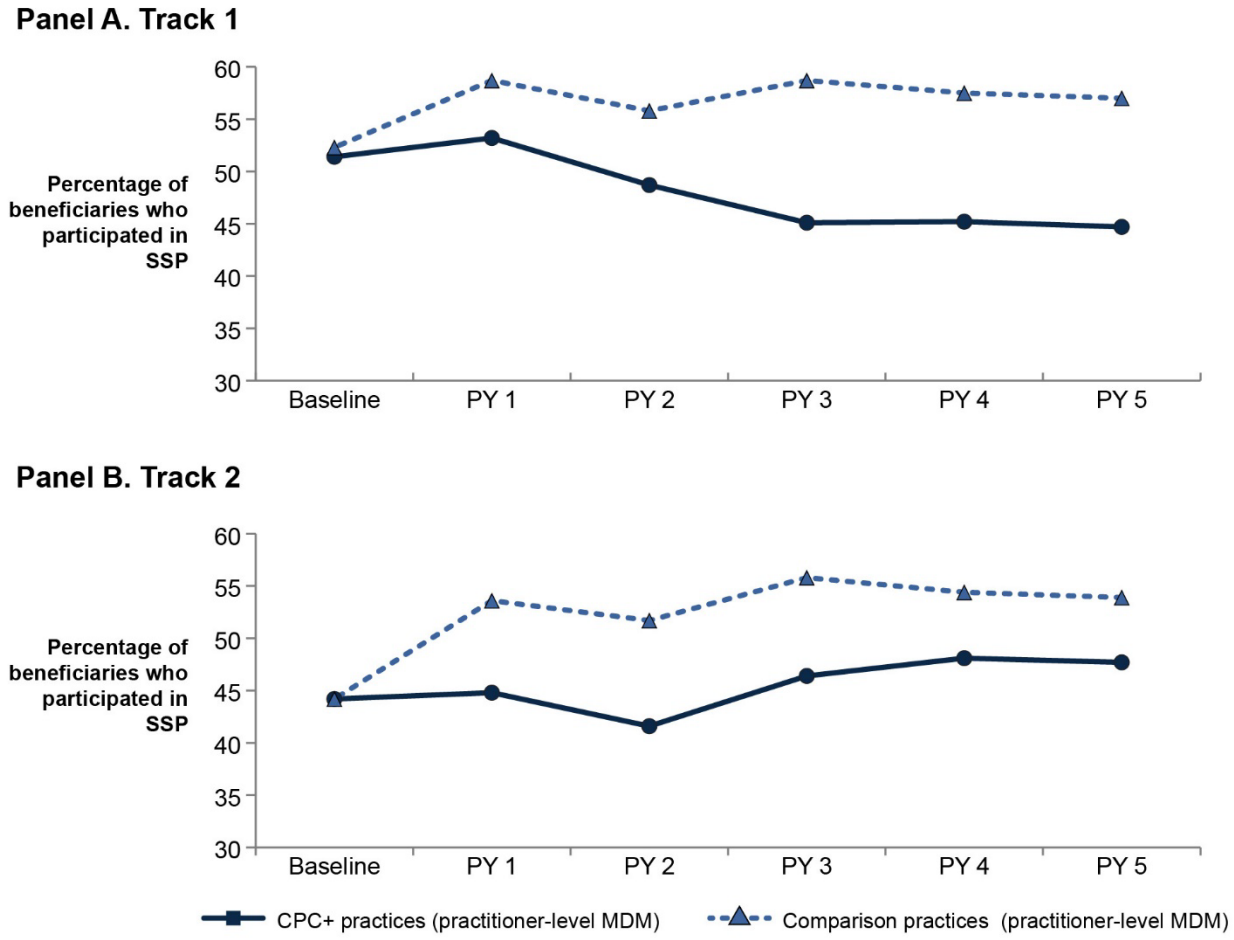
Participation in SSP was 42 to 59 percent and increased among comparison practices by 5 to 13 percentage points more than among CPC+ practices, depending on the program year and track. Figure 5.F.2 shows trends in participation in Medicare SSP based on the practitioner-level MDM. Participation in SSP among both CPC+ and comparison practices was large, with roughly half of the practices participating each year. Participation in SSP started off similar at baseline for CPC+ and comparison practices, with less than a 1 percentage point difference in participation for both Track 1 and Track 2 practices. From baseline to PY 5, participation in SSP among comparison practices increased, while among CPC+ practices the changes depended on track.

- For Track 1, participation in SSP among CPC+ practices declined by 6.7 percentage points between baseline and PY 5, while participation among comparison practices overall rose by 4.7 percentage points. Across the five program years, the difference in participation between CPC+ and comparison practices widened for Track 1, resulting in a -11.4 difference-in-differences estimate.
- For Track 2, we observed a slightly different pattern: from baseline to PY 5, participation among CPC+ practices increased by 3.5 percentage points, while participation among comparison practices increased by 9.7 percentage points, resulting in a -6.2 difference-in-differences estimate.

[107] For comparison selection, we measured baseline participation status for SSP and Next Gen as of January 1, 2017. Therefore, we measured participation in the first year of CPC+ as participation as of January 1, 2018, which was the end of PY 1; participation in the second year of CPC+ as participation as of January 1, 2019, which was the end of PY 2; participation in the third year of CPC+ as participation as of January 1, 2020, which was the end of PY 3; participation in the fourth year of CPC+ as participation as of January 1, 2021, which was the end of PY 4; and participation in the fifth year of CPC+ as participation as of December 31, 2021, which was the end of PY 5.

Figure 5.F.2. Trends of participation in SSP by beneficiaries in CPC+ practices and comparison practices in the baseline and five program years (based on practitioner-level MDM): Track 1 and Track 2

Beneficiary participation in SSP ranged from 42 to 59 percent. Participation decreased more among CPC+ practices than among comparison practices for Track 1 and grew less among CPC+ practices than among comparison practices for Track 2.



Source: Analysis of practitioner-level MDM extracts from February 27, 2017, February 23, 2018, February 26, 2019, February 28, 2020, February 26, 2021, and February 25, 2022.

Notes: We report participation in SSP as the percentage of beneficiaries who were exposed to SSP in each period in each track in each group (CPC+ or comparison practices), with comparison practices weighted using matching weights. We used the practitioner-level MDM to calculate participation as indicated in the chart legend in parentheses. We additionally weighted participation by the number of beneficiaries assigned to that practice during the baseline period, so that the results can be interpreted as the percentage of beneficiaries who were participating. The date used to define whether a practice participated in SSP at baseline was January 1, 2017 (consistent with the date used to define participation in comparison group selection). Accordingly, we defined the PY 1 participation value as participation as of January 1, 2018, the PY 2 participation value as participation as of January 1, 2019, the PY 3 participation value as participation as of January 1, 2020, and the PY 4 participation value as participation as of January 1, 2021. Because CPC+ ended in 2021, we defined participation for CPC+ PY 5 as participating as of December 31, 2021. In the practitioner MDM, 91 percent of participation in SSP is counted at the TIN level, while the remaining 9 percent is at the NPI-TIN level. If an NPI was listed in the practitioner MDM, we counted all practices with an NPI-TIN listed in that year as participating in SSP. If the NPI was missing in the practitioner MDM, we counted all practices with the TIN listed in that year as participating in SSP.

MDM = CMS Master Data Management System; NPI = National Provider Identifier; PY = Program Year; SSP = Medicare Shared Savings Program; TIN = Taxpayer Identification Number.

Among SSP participants, participation in tracks with downside risk increased from 10 to 17 percent to 35 to 52 percent (depending on track and CPC+ status). Figure 5.F.3 shows trends in participation in downside risk among SSP practices as described in the Methods section.[108] Participation in SSP tracks with downside risk started at similar levels at baseline (10 to 17 percent) for CPC+ and comparison practices in both tracks, with differences in participation less than 1 percentage point for Track 1 and less than 4 percentage point for Track 2.[109] From baseline to PY 5, participation in downside risk increased for both CPC+ and comparison practices; a large increase in participation occurred between PY 2 and PY 3, which might be explained by several new SSP track options (that is, Basic Tracks A, B, C, D, and E) becoming available in July 2019. The difference in participation rates in tracks with downside risk between CPC+ and comparison practices also changed over time, but the pattern of changes depended on track. Specifically:

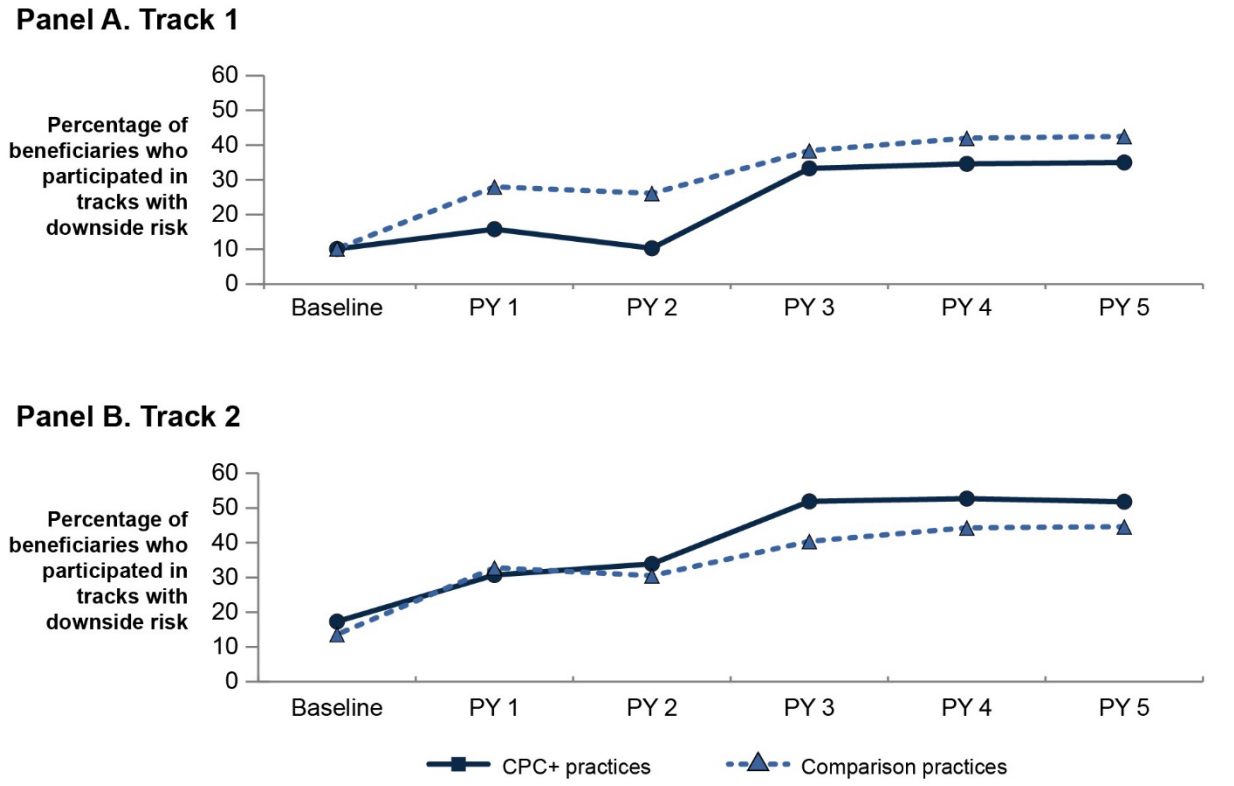
- For Track 1, participation grew more among comparison practices than among CPC+ practices. Between baseline and PY 5, participation increased by 25 percentage points among CPC+ practices and by 32 percentage points among comparison practices.
- For Track 2, participation grew more among CPC+ practices than among comparison practices. Participation among CPC+ practices increased by 35 percentage points between baseline and PY 5, while participation among comparison practices increased by 31 percentage points.

[108] We ran this analysis at the practice level because information about SSP tracks used to identify downside risk was only available at the practitioner level.

[109] This is consistent with the fact that we matched SSP practices on their baseline participation in the tracks with downside risk.

Figure 5.F.3. Trends of participation in SSP tracks with downside risk among SSP practices by beneficiaries in CPC+ practices and comparison practices in the baseline and five program years: Track 1 and Track 2

Beneficiary participation in SSP tracks with downside risk grew over time, with a large shift in growth between PY 2 and PY 3. Participation grew more among comparison practices (than among CPC+ practices) for Track 1 and grew more among CPC+ practices (than among comparison practices) for Track 2.



Source: Analysis of practitioner-level MDM extracts from February 27, 2017, February 23, 2018, February 26, 2019, February 28, 2020, February 26, 2021, and February 25, 2022.

Notes: We report participation in downside risk as the percentage of beneficiaries participating in SSP who were exposed to downside risk in each period in each track in each group (CPC+ or comparison practices), with comparison practices weighted using matching weights. We additionally weighted participation by the number of beneficiaries assigned to that practice during the baseline period, so that the results can be interpreted as the percentage of beneficiaries who were participating. The population we used to calculate participation was SSP practices, which varied by year. The date used to define whether a practice participated in SSP at baseline was January 1, 2017 (consistent with the date used to define participation in comparison group selection). Accordingly, we defined the PY 1 participation value as participation as of January 1, 2018; the PY 2 participation value as participation as of January 1, 2019; the PY 3 participation value as participation as of January 1, 2020; and the PY 4 participation value as participation as of January 1, 2021. Because CPC+ ended in 2021, we defined participation for CPC+ PY 5 as participating as of December 31, 2021. In the practitioner MDM, 91 percent of participation in SSP is counted at the TIN level, while the remaining 9 percent is at the NPI-TIN level. If an NPI was listed in the practitioner MDM, we counted all practices with an NPI-TIN listed in that year as participating in SSP. If the NPI was missing in the practitioner MDM, we counted all practices with the TIN listed in that year as participating in SSP. The numerator for participation includes practices that participated in SSP tracks with downside risk. The SSP tracks that involve downside risk include Track 1+, Track 2, Track 3, Basic C, Basic D, Basic E, and Enhanced; the SSP tracks that do not involve downside risk include Track 1, Basic A, and Basic B.

MDM = CMS Master Data Management System; NPI = National Provider Identifier; PY = Program Year; SSP = Medicare Shared Savings Program; TIN = Taxpayer Identification Number.

Participation in Next Gen remained lower than 5 percent but increased among comparison practices by 1 to 2 percentage points more than among CPC+ practices by PY 5. The CPC+ and comparison groups started out at close to 0 percent participation in the baseline period. This is because practices participating in CPC+ were not permitted to join Next Gen, and in the comparison selection process, we restricted potential comparison practices to those that were also not participating in Next Gen during the baseline period.[110] Participation among Track 1 CPC+ practices grew very little, to only 1.2 percent by PY 5 (because only CPC+ practices that stopped participating in CPC+ could join Next Gen); in contrast, participation among their comparison counterparts grew to 3.3 percent by PY 5. Track 2 experienced a very similar pattern: participation among CPC+ practices grew to 1.2 percent by PY 5, and participation among comparison group practices grew to 2.3 percent by PY 5. For Track 1 and Track 2, the PY 5 difference-in-differences estimates of -2.1 and -1.3 percentage points, respectively, are statistically significant at the 10 percent level.

Robustness checks using the beneficiary-level MDM showed lower levels of participation in SSP and Next Gen by up to 9.6 percentage points but similar trends in participation as the practitioner-level MDM, and thus similar difference-in-differences estimates (Table 5.F.2). See Appendix 5.D in the Fourth Annual Report for a discussion of the reason for the differences (Laird et al. 2022).

C. Participation in other primary care transformation initiatives

Participation in AHC was low among CPC+ and comparison practices. Rates of participation in AHC increased for both CPC+ and comparison groups from PY 2 (the first year of the model that beneficiaries were attributed) through PY 5, but participation remained less than 2 percent for both groups throughout the period.

Through PY 3, participation in TCPI fell among CPC+ practices and remained more constant among comparison practices.[111] TCPI participation among CPC+ practices remained stable between baseline and PY 1, and then decreased by 7.7 percentage points for Track 1 and 7.9 percentage points for Track 2 from PY 1 to PY 2. Participation then remained at the same small rate of about 2 to 3 percent in PY 3. Alternatively, comparison practices' participation remained relatively constant between 10 and 15 percent through PY 2, and then decreased in PY 3 to around 7 percent—possibly due to anticipation of the initiative's end.[112] This led to difference-in-differences estimates in the last year of TCPI (PY 3) of -4.5 percentage points and -2.5 percentage points for Track 1 and Track 2, respectively. The higher participation rate of the comparison group practices in TCPI suggests that some CPC+ practices would have participated in TCPI in the absence of CPC+. Since some comparison practices received learning supports under TCPI, the difference in learning supports between CPC+ and comparison practices is lower than the total learning supports that CPC+ practices receive through CPC+.

For both PCF and DC, participation rates were low among CPC+ and comparison practices. We found less than 3 percent participation in PCF and less than 4 percent participation in DC for CPC+ and

[110] Participation was not exactly zero, because the IQVIA practitioner rosters we use are not the same as the CMS rosters. Therefore, a couple of CPC+ practices are marked as participating in Next Gen based on the fact that at least one practitioner affiliated with the practice, according to the IQVIA data, had participated in Next Gen.

[111] We analyzed TCPI through PY 3 (i.e., 2019) because TCPI ended in September 2019.

[112] Although active CPC+ practices are not allowed to participate in TCPI, positive participation during CPC+ likely reflects additional belated withdrawals from TCPI, differences between the IQVIA roster of practitioners participating in CPC+ and the actual CMS CPC+ practitioner rosters, or practices that stopped participating in CPC+ (but are still included in the ITT sample) and joined TCPI.

comparison groups in PY 5 (the first performance year of the two models). CPC+ practices were technically unable to participate in PCF or DC during the CPC+ intervention period; however, we observed non-zero participation in PCF and DC among CPC+ practices, which may be explained by differences between the IQVIA and CMS practitioner rosters, or the ITT approach, which continues to follow practices that no longer participate in CPC+ or beneficiaries who are no longer attributed to a CPC+ practice.

D. Participation in CMS bundled payment initiatives

Participation in BPCI and BPCI Advanced was less than 2 percent and was similar for CPC+ and comparison practices. We found low levels of participation in the BPCI and BPCI Advanced initiatives for CPC+ and comparison groups in both tracks. For BPCI, which ended in September 2018, there were around 1 percentage point decreases in participation among both groups. For BPCI Advanced, which began in October 2018, for both tracks, there was a 0 to 0.1 percentage point increase in participation among CPC+ practices and 0.1 percentage point increase among comparison practices between PY 2 and PY 5. The lack of difference between CPC+ and comparison practices is not surprising, since BPCI and BPCI Advanced are national models and both CPC+ and comparison practices can participate in them.

E. Participation in population health initiatives

Participation in Million Hearts was less than 2 percent and was similar for CPC+ and comparison practices. For Million Hearts, which started in 2017, we observed 1 to 2 percent participation in PY 1 (the first year of the model that beneficiaries were attributed) for CPC+ and comparison groups in both tracks. Participation decreased to close to 0 in PY 2 through PY 4 (the most recent year of available data) for both practice groups in both tracks. Like BPCI and BPCI Advanced, Million Hearts is a national model that permits both CPC+ and comparison practices to participate, which might explain the similar participation levels between CPC+ and comparison practices.

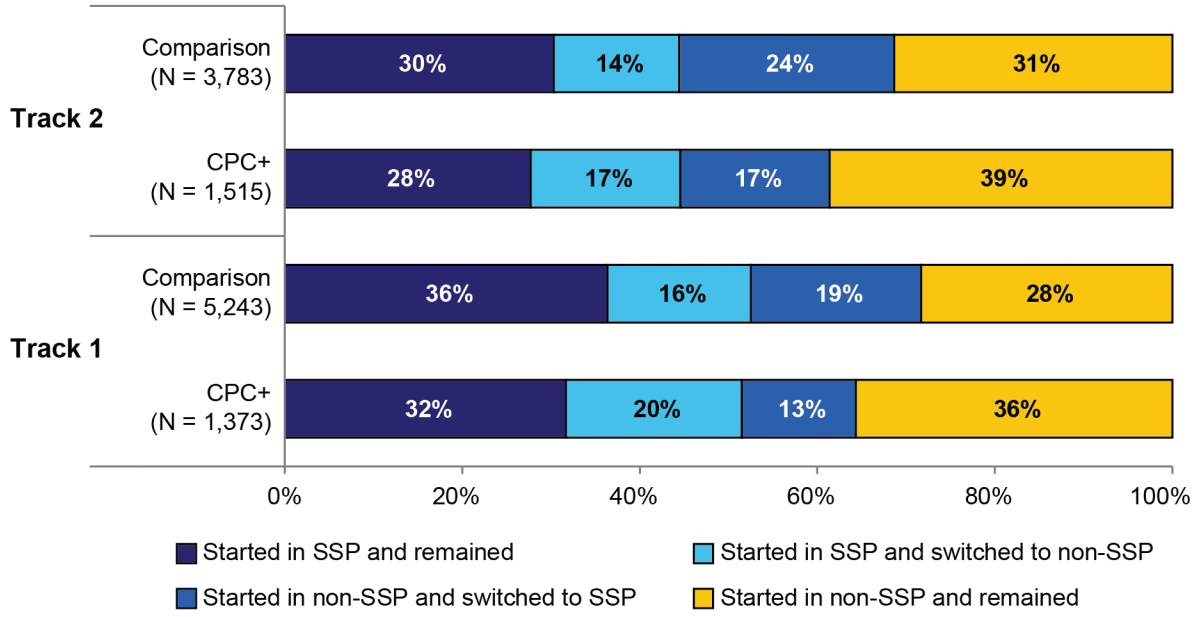
5.F.4. Implications for CPC+ impact analyses

The moderately larger increases in participation in Medicare FFS value-based purchasing models and smaller decreases in participation in TCPI for comparison group practices compared to CPC+ practices could decrease the marginal impact of the CPC+ incentives and supports in improving primary care, relative to a case in which these other initiatives did not exist. That is, if these other initiatives are encouraging care delivery changes in the comparison group similar to those occurring in the CPC+ group, and the changes improve outcomes, we may observe only small effects of CPC+ or none at all, even if the CPC+ model of care transformation is indeed effective in improving quality or lowering costs. However, the initiative for which these differential changes in participation between the CPC+ and comparison group is the largest—SSP—is a nationwide program, and the comparison group's participation likely represents the correct counterfactual to the scenario in which CPC+ did not exist. Based on findings from our analysis, we will conduct a sensitivity test to our primary impact analyses, controlling for contemporaneous SSP. This will shed light on whether our impact results are at all driven by SSP participation.

Given the sizeable differential changes in participation between the CPC+ and comparison groups in SSP during the intervention period, the findings from the impact evaluation for the SSP subgroups, which are defined based on SSP status at baseline, should be interpreted with caution. In fact, as illustrated in Figure 5.F.4, about one-third of the practices switched in or out of SSP from baseline through PY 5. Among

practices that changed SSP status, more practices switched into SSP in the comparison group, whereas in the CPC+ group, more practices switched out of SSP, resulting in increasing participation in SSP over time among comparison practices relative to CPC+ practices. Hence, instead of interpreting the SSP subgroup estimates as the impact of CPC+ combined with SSP throughout the intervention period, these estimates should be interpreted as the impact of *starting* CPC+ in SSP.

Figure 5.F.4. Distribution of SSP switching patterns in CPC+ practices and comparison practices: Track 1 and Track 2



Source: Analysis of practitioner-level MDM extracts from February 27, 2017, February 23, 2018, February 26, 2019, February 28, 2020, February 26, 2021, and February 25, 2022.

Notes: In each track in each group (CPC+ or comparison practices), each practice fell into one of the four categories of SSP switching patterns from baseline through the end of CPC+: (1) "started in SSP and remained," meaning that practices started CPC+ while participating in SSP and remained in SSP throughout the intervention period (shown in dark blue bars); (2) "started in SSP and switched to non-SSP," meaning that practices started CPC+ while participating in SSP but switched to non-SSP in any PY during the intervention period (shown in light blue bars); (3) "started in non-SSP and switched to SSP," meaning that practices started CPC+ in non-SSP but switched to SSP in any PY during the intervention period (shown in blue bars); or (4) "started in non-SSP and remained," meaning that practices started CPC+ in non-SSP and remained in non-SSP throughout the intervention period (shown in yellow bars). The date used to define whether a practice participated in SSP at baseline was January 1, 2017 (consistent with the date used to define participation in comparison group selection). Accordingly, we defined the PY 1 participation value as participation as of January 1, 2018, the PY 2 participation value as participation as of January 1, 2019, the PY 3 participation value as participation as of January 1, 2020, and the PY 4 participation value as participation as of January 1, 2021. Because CPC+ ended in 2021, we defined participation for CPC+ PY 5 as participating as of December 31, 2021. In the practitioner MDM, 91 percent of participation in SSP is counted at the TIN level, while the remaining 9 percent is at the NPI-TIN level. If an NPI was listed in the practitioner MDM, we counted all practices with an NPI-TIN listed in that year as participating in SSP. If the NPI was missing in the practitioner MDM, we counted all practices with the TIN listed in that year as participating in SSP.

MDM = CMS Master Data Management System; NPI = National Provider Identifier; PY = Program Year; SSP = Medicare Shared Savings Program; TIN = Taxpayer Identification Number.

5.G. Estimated impacts of Comprehensive Primary Care Plus on potential opioid overuse

In this Appendix, we examine the impact of the Comprehensive Primary Care Plus (CPC+) model on the potential overuse of prescription opioids for Medicare fee-for-service (FFS) beneficiaries during the five years of CPC+. We also assess differential changes in the prescribing behaviors of clinicians providing care as part of CPC+ and comparison practices. In Section 5.G.1, we describe the motivation for this analysis, including an overview of how CPC+ could affect potential opioid overuse. We next explain the analytic methods, study population, and key outcomes of interest (Section 5.G.2). Finally, we describe the results (Section 5.G.3) and discuss their implications and the limitations of this analysis (Section 5.G.4).

5.G.1. Introduction

Using opioids at high doses carries substantial risks of serious adverse effects, including addiction, misuse, serious fractures, cardiovascular events, and overdose, especially if administered long term (Chou et al. 2020; Els et al. 2017; Von Korff et al. 2011;). Older adults have a particularly high risk of adverse events and mortality related to high-dose opioid use due to comorbidities and polypharmacy (Lehmann and Fingerhood 2018; Song 2017, Yoshikawa et al. 2020). Although opioid dispensing rates per capita have been decreasing since 2013 (CDC 2019), reducing opioid misuse remains a goal for Medicare beneficiaries (CMS 2020). Concrete steps toward that goal include some combination of reducing initial opioid prescriptions, avoiding increasing dosages, careful tapering, and treating opioid use disorder while pursuing alternative treatments for pain (CMS 2020; CRS 2021, Humphreys and Marsden 2022).

Medicare beneficiaries often use opioids for chronic pain and typically obtain them from a physician rather than other sources (Schepis et al. 2020). Further, primary care practitioners write the largest volume of Medicare Part D opioid prescriptions (Chen et al. 2016). Therefore, interventions that focus on prescribing in primary care have potential to reduce harm from opioid overuse among Medicare beneficiaries. Several mechanisms could contribute to the success of such interventions. For example, integrating behavioral health staff and clinical pharmacists into the primary care setting (Seal et al. 2020) and enhancing medication management practices (Parchman et al. 2019) may reduce high-dose opioid prescribing in primary care practices. Improving care management processes and providing clinician education have also resulted in improved adherence to opioid prescribing guidelines (Liebschutz et al. 2017; Meisenberg et al. 2018). However, past studies have been limited in the number of practice sites and patients.

In 2017, the Centers for Medicare & Medicaid Services (CMS) launched CPC+, which included 2,905 primary care practices in 18 regions of the United States. CPC+ was a five-year model aimed at improving quality of care and reducing costs. Primary care practices that participated in CPC+ received enhanced Medicare payments to support five functions: access and continuity of care, care management, comprehensiveness and coordination, patient and caregiver engagement, and planned care and population health.

CPC+ did not explicitly target opioid prescribing, but the participating practices were required to implement several care delivery approaches that could change prescribing behaviors and increase attention to potential opioid overuse. Those approaches included delivering comprehensive medication management (CMM), by clinical staff within the practice or a pharmacist, screening for mental health and substance use disorders, and either co-locating a credentialed behavioral health staff member in the practice or designating a practitioner or team member to provide care management for behavioral health

conditions. Practices chose to join one of two CPC+ tracks. Practices in both tracks received enhanced Medicare payment for participating in CPC+ and for improving their performance on cost, utilization, and/or quality measures. Relative to Track 1 practices, those in Track 2 received larger Medicare payments to support more enhanced care delivery; for example, Track 2 practices were required to implement CMM for all patients who received longitudinal care management (those with complex needs and those who experienced transitions of care) (Peikes et al. 2021b).

This analysis examines whether CPC+ impacted potential opioid overuse among Medicare FFS beneficiaries who use opioids long term. We compare changes over time in potential opioid overuse among beneficiaries assigned to CPC+ practices versus comparison practices. We also assess differential changes in the prescribing behaviors of clinicians providing care as part of CPC+ and comparison practices. Our analysis seeks to understand whether a large-scale practice transformation that was not explicitly designed to impact the use of opioids but included some of the core components of previous opioid use interventions could lead to reductions in opioid overuse. Because opioid overuse carries a risk of overdose, which greatly accelerated during the COVID-19 pandemic in the general population (AMA 2022a), we also descriptively assess trends in opioid overdose.

5.G.2. Methods

A. Study design and setting

Using a difference-in-differences framework, we estimated the impacts of CPC+ on potential opioid overuse using Medicare FFS claims and Medicare Part D prescription drug event data over the baseline period (2016) and five program years of CPC+ (2017 through 2021). The analysis included 1,373 Track 1 practices, 1,515 Track 2 practices, and 6,921 comparison practices. Following the same approach as for the main CPC+ impact analysis, to create the analytic sample, we (1) attributed beneficiaries to all CPC+ and potential comparison practices; (2) applied an intent-to-treat (ITT) approach for practices and beneficiaries, and (3) selected matched comparison practices.

Every quarter, we attributed beneficiaries to the practice that delivered the largest share of their primary care visits over the prior two years. Following an ITT approach, we retained practices in the analysis regardless of whether they disenrolled from CPC+. We then applied an ITT approach for beneficiaries by assigning beneficiaries to the first practice to which they were attributed during the baseline period and continuing to assign the beneficiary to the same practice throughout the baseline period regardless of whether the beneficiary continued to receive care at that practice. We repeated the same process for the intervention period, assigning beneficiaries to the first practice to which they were attributed after the intervention began. This ITT approach helps to avoid the potential biases on impact estimates that could arise if we analyzed only the beneficiaries who remained attributed to practices over time or the practices that remained in the sample. Our sample was therefore a repeated cross-section of beneficiaries with a high degree of overlap in the sample across program years (Peikes et al. 2021b).

We relied on the same external comparison group used for the main impact analysis; that is, we used matched comparison practices drawn from areas located near the CPC+ regions but often out of state (Ghosh et al. 2020; Kranker et al. 2021). Although this approach ensures that the CPC+ and comparison groups are well balanced at baseline, large changes in sample composition could confound impact estimates. Therefore, using baseline characteristics, we examined changes in sample composition over time, and we also controlled for baseline beneficiary characteristics in all our regressions.

B. Outcomes

B.1. Potential opioid overuse

We defined potential opioid overuse as opioid use at a daily dosage of 90 morphine milligram equivalents (MMEs) or more among beneficiaries who use opioids long term. Long-term use means having an opioid supply of 90 or more days in a calendar year with no more than a 7-day gap between prescriptions. We analyzed data at a beneficiary-year level, with the overuse measure equal to one for beneficiaries with overuse and equal to zero for those without overuse in a year. This definition follows the specifications for the Electronic Clinical Quality Improvement Measure (eCQM) 460 (eCQI Resource Center n.d.). The main difference between our potential opioid overuse measure and the eCQM 460 is that our measure relies on Part D claims data, whereas the eCQM 460 relies on electronic health record (EHR) data. The key advantages of using Part D claims data are (1) claims data capture prescription fills, not just prescribing behavior; (2) dosage information is more accurate in claims data; and (3) claims data capture filled prescriptions from all prescribers.

The 90 MME threshold for high dosage aligns with the 2016 Centers for Disease Control and Prevention (CDC) prescribing guidelines that asked clinicians to avoid increasing dosages beyond that threshold (Dowell et al. 2016). Although the 2022 guidelines removed the 90 MME threshold to address concerns about sudden discontinuation and abrupt tapering, they continue to warn prescribers about the risks of high-dosage use and encourage them to carefully evaluate risks and benefits of dosage increases (Dowell et al. 2022). Concerns about discontinuation and abrupt tapering are not relevant to our analysis because CPC+ did not use this measure as part of the intervention nor assess practices' performance during the initiative.

To be included in the overuse measure, a beneficiary had to (1) be assigned to a practice and (2) be continuously enrolled in Medicare Parts A, B, and D throughout each calendar year or until death. Following the eCQM specifications, we excluded beneficiaries for whom opioid use was likely appropriate: beneficiaries with a diagnosis of cancer during or one year before the measurement year and those with a diagnosis of sickle cell disease or with hospice use during the measurement year. To identify diagnoses for exclusion criteria, we used ICD-10 codes from the eCQM specifications. Because the eCQM specifications do not list national drug codes (NDCs), we relied on NDCs for opioid therapy from the Medication List Directory value sets for the Healthcare Effectiveness Data and Information Set (HEDIS®) measure Use of Opioids at a High Dosage to identify beneficiaries who used opioid therapy (National Committee for Quality Assurance 2020, 2021). We used the CDC Opioid NDC and Oral MME Conversion File (CDC 2021) to calculate daily MME for beneficiaries with opioid therapy use.

B.2. Long-term opioid use

We defined long-term use as receipt of opioid therapy for 90 days or more during the measurement year, with no more than a 7-day gap between prescriptions. To be included in the long-term use measure, a beneficiary had to (1) be assigned to a practice; (2) be continuously enrolled in Medicare Part A, B, and D throughout each calendar year or until death; and (3) have at least one opioid prescription during the measurement year (that is, had to have some opioid use).

Because long-term use is the denominator for the potential opioid overuse measure, impacts on long-term use could make it difficult to interpret the estimated impacts on overuse. Namely, if the comparison group reduced long-term use more than the CPC+ group did, it might be more difficult for the comparison group to reduce overuse among the remaining beneficiaries who use opioids long term, if they are of higher risk

than beneficiaries who no longer use opioids long term. In other words, performing better on the long-term use measure might make it more difficult to perform better on the overuse measure. To improve our understanding of this potential issue, we estimated the impact of CPC+ on long-term opioid use as a supporting analysis.

B.3. Prescribing

We compared several aspects of prescribing for CPC+ and comparison prescribers over time: (1) total MMEs per prescription (total dosage over all days of supply), (2) average number of days' supply per prescription (duration), and (3) number of prescriptions per year (volume). We analyzed dosage and duration per prescription, rather than in total, to account for potential impacts of CPC+ on the number of opioid prescriptions. Our choice of prescribing outcomes was consistent with the literature on opioid prescribing (Cramer et al. 2021; Delgado et al. 2018; Duan et al. 2022; Keshwani et al. 2022; March et al. 2022; Schommer et al. 2020).

B.4. Opioid overdose

Because opioid overuse carries the risk of overdose, we descriptively analyzed changes in overdoses over time for beneficiaries assigned to CPC+ and comparison practices. We identified opioid overdoses based on diagnosis codes for overdose in inpatient, outpatient, or carrier claims using ICD-10 codes from Lo-Ciganic et al. 2019. Opioids were classified as synthetic opioids, including tramadol/fentanyl (T40.4); heroin (T40.1); prescription opioids (T40.2–40.3); and other opioids, including opium (T40.0) and other/unspecified narcotics (T40.6). We identified opioid overdoses among beneficiaries who use opioids long term.

C. Statistical analysis

C.1. Potential opioid overuse and long-term opioid use

We estimated difference-in-differences models separately by track, reflecting the differences in care delivery requirements and payments. We tested whether estimated impacts differ based on disability status and dual eligibility for Medicare and Medicaid because these populations are especially vulnerable to opioid misuse (Buchmueller and Carey 2018). We also examined whether impacts differed among practices that participated in the Medicare Shared Savings Program (SSP) at baseline and those that did not because estimated impacts might be larger among practices that participate in both CPC+ and SSP.

To estimate the cumulative impact of CPC+ on potential opioid overuse and long-term opioid use, we used a single intervention indicator for the five years combined. The year immediately preceding the start of CPC+ (2016) was the reference category, or the baseline year, for obtaining the difference-in-differences impact estimates; that is, the impact estimate was the CPC+ and comparison difference in an outcome in the intervention period minus the average CPC+ and comparison difference in the baseline year. We estimated separate regression models for each outcome of interest. Our main estimation approach is shown in Equation 5.G.1:

$$(5.G.1) \qquad Y_{ijt} = \alpha + \beta X_{it} + \delta C_j \cdot Year_{t=2020,2021} + \tau S_{jt} + \rho M_{jt} + \gamma P_j + \gamma Year_t + \theta Treatment_j \cdot Post_t + \varepsilon_{ijt},$$

where

Y = outcome variable for beneficiary i, in practice j, in year t.

The control variables included the following:

- X = vector of beneficiary-level controls measured at the start of the baseline period ($t = 2016$) and the start of the intervention period ($t = 2017$); see section below on control variables.
- C = vector of control variables that measure the intensity of the COVID-19 epidemic and the government response, which are interacted with year indicators for 2020 and 2021.
- S = two sets of variables that address state opioid policies: (1) sophistication of prescription drug monitoring programs (PDMPs) for all years except 2017 and (2) value of federal opioid funding awarded to each state for all years. See section below on control variables.
- M = county-level pharmaceutical to physician marketing intensity of opioids for all years.
- P = fixed effects for practice j that controls for all time-invariant practice characteristics.
- *Year* = binary indicator for each year t in the intervention period.
- *Treatment* = binary indicator of treatment status, that is, of being attributed to a CPC+ practice.
- *Post* = binary indicator for the intervention period.
- ε = idiosyncratic error term.

The difference-in-differences impact estimate over the five-year intervention period is θ. Note that the treatment indicator is collinear with the practice fixed effects; therefore, the difference between CPC+ and comparison practices at baseline, or the main effect of treatment status, cannot be estimated. However, the model can estimate the interaction term between treatment and Post, that is, the difference-in-differences estimate over the entire intervention period. We also estimated a similar model to obtain impact estimates for each intervention year by interacting the treatment dummy with each intervention year.

Given large sample sizes, we reduced computational intensity by relying on linear models; further, nonlinear models can have challenges with convergence. Prior studies of similar interventions showed that linear models provided similar results as nonlinear models (Laird et al. 2022).

To test whether any estimated impacts might be due to pre-intervention differences in outcome trends between CPC+ and comparison groups, we performed falsification tests by estimating "impacts" of CPC+ in 2016 (the year before CPC+ started), with 2015 as the baseline year. For this test, we constructed outcomes for an additional year (2015); however, data were not available to construct outcomes for earlier years.

We report two-sided *p*-values, which we consider statistically significant if $p < 0.05$. However, we did not rely entirely on *p*-values; instead, to avoid spurious conclusions, we examined the consistency of the magnitude and statistical significance of the estimates and the patterns of findings across time periods, tracks, and outcomes. We also considered information about model implementation. This approach is supported by the American Statistical Association's statement on statistical significance, which recommends that policy decisions should not be based entirely on whether a *p*-value passes a specific threshold and argues that *p*-values cannot entirely measure the importance of a result (Wasserstein and Lazar 2016). Because we relied on the patterns of findings, rather than statistical significance for any one given finding, we did not adjust *p*-values for multiple comparisons. Standard errors were adjusted for clustering at the practice level, accounting for correlation in an outcome across beneficiaries assigned to the same practice—both within and across time periods.

Control variables. Each regression controlled for beneficiary characteristics, practice fixed effects, changes in state-level opioid policies and county-level opioid marketing intensity, and COVID-19 controls.

- **Beneficiary characteristics.** For observations in the intervention period, we measured beneficiary-level control variables directly before the start of CPC+ (that is, based on 2016 data). For observations in the baseline period, we measured beneficiary-level control variables directly before the start of the yearlong baseline period (based on 2015 data). In addition to the Hierarchical Condition Category (HCC) score, for comprehensive risk adjustment, the regression additionally includes indicators for specific chronic conditions selected based on their weight in the HCC score calculation and on a high prevalence in the CPC+ sample (collapsing categories, where appropriate). To account for possible changes in the relationship between characteristics measured at the start of the intervention and outcomes, we also included interactions between the HCC score and each intervention year from the second year onward as well as interactions between specific chronic conditions and the intervention year (Table 5.G.1).

Table 5.G.1. Beneficiary-level control variables for the difference-in-differences regressions

Baseline characteristic category	Variables[a]
Demographics	Age categories < 65 65–74 (reference category) 75–84 ≥ 85 Race/ethnicity[b] Non-Hispanic White (reference category) Black Hispanic All other/unknown Gender (binary indicator for male)
Original reason for Medicare eligibility	Original Medicare eligibility categories Age (reference category) Disability only ESRD only or ESRD with disability
Dual eligibility	Indicator for dual status (where dual is defined as those with full or partial Medicaid benefits according to Master Beneficiary Summary File)
Chronic conditions	HCCs:[c] HCC 18 – Diabetes with Chronic Complications HCC 19 – Diabetes without Complications HCC 21 – Protein-Calorie Malnutrition HCC 22 – Morbid Obesity HCC 23 – Other Significant Endocrine and Metabolic Disorders HCC 84 – Cardio-Respiratory Failure and Shock HCC 85 – Congestive Heart Failure HCC 96 – Specified Heart Arrhythmias HCC 106 – Atherosclerosis of the Extremities with Ulceration or Gangrene HCC 111 – Chronic Obstructive Pulmonary Disease HCC 135 – Acute Renal Failure HCC 138 – Chronic Kidney Disease, Moderate (Stage 3) HCC 173 – Traumatic Amputations and Complications HCC 40 or 47 – Rheumatoid Arthritis and Inflammatory Connective Tissue Disease or Disorders of Immunity HCC 46 or 48 – Severe Hematological Disorders, or Coagulation Defects and Other Specified Hematological Disorders HCC 57 or 58 – Schizophrenia or Major Depressive, Bipolar, and Paranoid Disorders HCC 70 or 71 – Quadriplegia or Paraplegia HCC 80 or 82 – Coma, Brain Compression/Anoxic Damage or Respirator Dependence/Tracheostomy Status HCC 86, 87, or 88 – Acute Myocardial Infarction, Unstable Angina and Other Acute Ischemic Heart Disease, or Angina Pectoris HCC 99 or 100 – Cerebral Hemorrhage, or Ischemic or Unspecified Stroke HCC 107 or 108 – Vascular Disease, with Complications HCC 157 or 158 – Pressure Ulcer of Skin with Necrosis Through to Muscle, Tendon, or Bone; or of Skin with Full Thickness Skin Loss HCC 8, 9, 10, 11, or 12 – Metastatic Cancer and Acute Leukemia; Lung and Other Severe Cancers; Lymphoma and Other Cancers; Colorectal, Bladder, and Other Cancer; or Breast, Prostate, and Other Cancers and Tumors Chronic Conditions Warehouse indicator: Alzheimer's disease or dementia Chronic condition indicators interacted with follow-up year from second follow-up year onward

Table 5.G.1. ***(continued)***

Baseline characteristic category	Variables[a]
Risk score	HCC score Indicator for whether HCC score was assigned a new enrollee HCC score (that is, HCC score was calculated based on demographic characteristics only) HCC score interacted with follow-up year from second follow-up year onward Indicator for being assigned a new enrollee HCC score interacted with follow-up year from second follow-up year onward

[a] We measured beneficiary-level control variables either directly before the start of CPC+ (for the intervention period observations) or directly before the start of the yearlong baseline period (for the baseline-period observations). The yearlong baseline period is 2016.

[b] We controlled for race/ethnicity with imputed race and ethnicity data, using a methodology called Medicare Bayesian Improved Surname Geocoding (MBISG 2.1). The set of MBISG race/ethnicity variables included imputed probabilities that each beneficiary is White, Black, Hispanic, or other. These probabilities, which incorporated administrative data, surname, and residential location, are strongly predictive of self-reported race and ethnicity (Haas et al. 2019).

[c] We selected a subset of the 79 HCCs—created by the HCC model for inclusion as control variables, based on the relative weight of specific HCCs in HCC score calculation as well as their prevalence in our analysis sample. We also included an indicator for Alzheimer's disease or dementia from the Chronic Conditions Warehouse (because there is not an HCC for Alzheimer's disease or dementia).

CPC+ = Comprehensive Primary Care Plus Initiative; ESRD = end-stage renal disease; HCC = Hierarchical Condition Category.

- **Practice fixed effects.** The practice fixed effects are indicators or dummy variables—one for each practice in the CPC+ and comparison groups. These account for any inherent, time-invariant differences between the CPC+ and comparison practices, whether such differences are observed or unobserved. Including practice fixed effects ensured that we accounted for any remaining imbalance in the practice-level variables used in matching and in any other unmeasured practice characteristics at baseline, when obtaining the difference-in-differences impact estimates. We did not incorporate changes over time in observed practice characteristics as control variables because CPC+ might have affected practice characteristics.
- **Changes in opioid-related regional controls.** Because comparison practices were generally located in different states than CPC+ practices, the estimated effect of CPC+ on potential opioid overuse could be due to differential changes in regional policies. To account for this potential confounding, the regression models controlled for the following:
  - Sophistication of PDMPs because using PDMPs has been found to be associated with reductions in opioid prescribing (Buchmueller and Carey 2018; Pardo 2017; Wen et al. 2017). We measured PDMP characteristics in all years except 2017; data for 2017 were not available (TTAC 2021). We created a simple index of PDMP characteristics, listed in Table 5.G.2, giving each characteristic an equal weight.
  - The amount of federal opioid grant funding per capita awarded to states in 2016 through 2021, to control for differential resources by state (Table 5.G.2). These grants provide funding for opioid programs and interventions as well as interventions to address substance use disorder with an opioid-related component. Grants that span several years were attributed to the first year of the award (Katcher and Ruhm 2021). We collected data directly from the funding agencies' data repositories, which include Substance Abuse and Mental Health Services Administration (SAMHSA), National Institutes of Health (NIH), CDC, Health Resources and Services Administration (HRSA), and Department of Justice (DOJ) (which captures funding from other agencies including Bureau of Justice Assistance [BJA], Office of Juvenile Justice and

Delinquency Prevention [OJJDP], and Office for Victims of Crime [OVC]). We divided total opioid funding for each state and year by census population data for each state and year. We coded state funding variables as zero for the years with no funding listed from the agencies from which we collected data. We used several specifications (and two different sources of funding data) for regional variables and found that results did not differ.

– The intensity of pharmaceutical opioid marketing to physicians, which has been found to increase prescribing (Hadland et al. 2019; Hollander et al. 2020). We measured the number of pharmaceutical direct-to-physician marketing events of opioids per clinician in a county in each year from 2016 through 2021. We created a measure of opioid marketing intensity using the CMS Open Payments database, which contains publicly accessible information about payments and transfers of value that reporting entities (pharmaceutical companies) make to covered recipients (physicians). We identified a marketing event involving opioids by using the same set of NDCs used to create the long-term use of opioids and potential opioid overuse outcomes.

Table 5.G.2. State- and county-level opioid policy controls for the difference-in-differences regressions

State opioid policy category	Components/description	Data source
Index for sophistication of PDMPs	Simple index (each component receives an equal weight) measured in 2016, 2018, 2019, 2020, and 2021 using the following components: • Uses a vendor • Integrated with EHRs • Integrated with pharmacy databases • Enrollment is mandatory for prescribers • Use is mandatory for prescribers • Enrollment is mandatory for dispensers • Use is mandatory for dispensers • Reports available to prescribers • Data collection frequency: real-time, daily, or next business day • Data sharing: share data with 21 states or more	PDMP TTAC 2022
State-level federal opioid funding	Amount of federal grants awarded in fiscal years 2016, 2017, 2018, 2019, 2020, and 2021	SAMHSA, NIH, CDC, HRSA, DOJ
Opioid marketing intensity	Number of pharmaceutical company direct-to-physician marketing events for opioids per clinician in a county in 2016, 2017, 2018, 2019, 2020, and 2021	CMS Open Payments

CDC = Centers for Disease Control and Prevention; CMS = Centers for Medicare & Medicaid Services; DOJ =Department of Justice; EHRs = electronic health records; HRSA = Health Resources and Services Administration; NIH = National Institutes of Health; PDMP = prescription drug monitoring program; SAMHSA = Substance Abuse and Mental Health Services Administration; TTAC = Training and Technical Assistance Center.

- **COVID-19 controls.** Because the intensity and effects of COVID-19 varied by region over time, we also controlled for several measures intended to capture differences in the effect of COVID-19 in CPC+ and comparison regions in 2020 and 2021. First, we used a state–hospital referral region (HRR)-level measure of excess deaths for all Medicare FFS beneficiaries in the CPC+ and comparison regions for each wave of the pandemic in 2020 and 2021. (We defined waves as follows: wave 1: March–May 2020, wave 2: June–September 2020, wave 3a: October–December 2020, wave 3b: January–February 2021, wave 4: March–May 2021, wave 5: June–October 2021, and wave 6: November–December 2021.) We created the excess deaths measure by following the methods in

Polyakova et al. (2021). Second, we used a publicly available social vulnerability index at the census tract level.[113] Third, we used a publicly available pandemic vulnerability index at the county level, calculated for each wave.[114] And fourth, we used a publicly available state government response index at the state-year level.[115]

Weighting. We applied weights to the observations in the regressions to ensure that the CPC+ and comparison groups were comparable. The regression weights equaled the covariate-balancing propensity score–based weights used to balance the CPC+ and comparison practices on their baseline characteristics. As is typical for propensity score weighting, we set the weights for the intervention practices at one, meaning that each intervention practice would count equally in practice-level analysis and each intervention beneficiary would count equally in beneficiary-level analysis. To achieve better balance between the intervention and comparison practices, the comparison practice weights varied based on the practice's similarity to the intervention group practices (Kranker et al. 2021).

Falsification tests. To test whether any estimated impacts might be due to pre-intervention differences in outcome trends between CPC+ and comparison groups, we performed falsification tests by estimating "impacts" of CPC+ in 2016 (the year before CPC+ started), with 2015 as the baseline year. For this test, we constructed outcomes for an additional year (2015); however, data were not available to construct outcomes for earlier years.

C.2. Prescribing

Beneficiaries attributed to CPC+ or comparison practices could have obtained opioid prescriptions from clinicians in their attributed practices or from outside prescribers. To understand which group of clinicians reduced their prescribing the most, we examined changes in prescribing to beneficiaries who use opioids long term, by clinicians affiliated with CPC+ practices (CPC+ prescribers), clinicians affiliated with comparison practices (comparison prescribers), and clinicians not affiliated with either CPC+ or comparison practices who also delivered care to beneficiaries in our sample (outside prescribers). We first checked whether outside prescribing to CPC+ beneficiaries and comparison beneficiaries changed significantly.

Next, we estimated the impact of CPC+ on the prescribing outcomes by assessing whether CPC+ prescribers had greater improvements in these outcomes than comparison prescribers. The study population consisted of CPC+ or comparison clinicians (as defined by the National Provider Identifier [NPI]), who prescribed opioids to the beneficiaries included in the analyses of potential opioid overuse. The unit of observation in this analysis is the clinician-year. If a clinician was affiliated with multiple practices (within our sample of CPC+ and comparison practices) in a given year, we randomly assigned that clinician to a single practice. Approximately 3 to 4 percent of the clinicians were affiliated with multiple practices each year, depending on sample (CPC+ or comparison), track and year.

To identify the practice with which an NPI was affiliated, we purchased yearly rosters from IQVIA, a commercial health care data vendor that maintains and verifies lists of practitioners who work in practices throughout the country. We then used data from IQVIA's SK&A database for the baseline period and the first two years of CPC+, and data from IQVIA's OneKey database starting in Program Year (PY) 3,

[113] See https://www.atsdr.cdc.gov/placeandhealth/svi/index.html.

[114] See https://www.niehs.nih.gov/research/programs/coronavirus/covid19pvi/details/index.cfm.

[115] See https://www.bsg.ox.ac.uk/research/research-projects/covid-19-government-response-tracker.

because IQVIA discontinued the SK&A data and replaced it with OneKey data in 2019. To understand the effect of the source data change from SK&A to OneKey, we conducted a sensitivity test where we limited the sample of NPIs to those who were affiliated with CPC+ or comparison practices at baseline or at the start of the intervention period (thereby not allowing the addition of new NPIs from PY 2 onward). The results remained qualitatively the same with and without the additional sample restriction (data available upon request).

As in the potential overuse analysis, we used difference-in-differences models with practice-level matching weights and used the same baseline and follow-up periods. We controlled for the prescriber's age, gender, and an indicator for whether the prescriber was a physician (versus a nurse practitioner or a physician's assistant), because these characteristics influence prescribing behaviors (Baker et al. 2022; Tamblyn et al. 2022). We clustered standard errors at the practice level and included practice fixed effects, and all other regional control variables as in the potential opioid overuse analysis. We also ran a falsification test for prescribing outcomes.

C.3. Opioid overdose

We analyzed trends in the percent of beneficiaries with any opioid overdose and overdose from prescription opioids, separately for CPC+ and comparison beneficiaries and by track. We did not estimate the impact of CPC+ on opioid overdoses because of insufficient power to detect effects.

5.G.3. Results

A. Beneficiary sample for potential opioid overuse

A.1. Inclusion criteria

After we applied the inclusion criteria, at baseline, 40,219 beneficiaries with long-term opioid use were attributed to Track 1 CPC+ practices (6.6 percent of all continuously enrolled FFS beneficiaries with Parts A, B, and D); among them, 7,743 beneficiaries (19.3 percent) had potential opioid overuse. In comparison practices, 129,178 beneficiaries were using opioids long term at baseline (6.5 percent of continuously enrolled FFS beneficiaries with Parts A, B, and D); among them, 24,285 beneficiaries (18.8 percent) had potential opioid overuse. We retained similar proportions of beneficiaries after each inclusion criteria, regardless of track or CPC+ status; that is, we found no evidence that CPC+ and comparison beneficiaries in either track differed on inclusion criteria (Tables 5.G.3 and 5.G.4).

Table 5.G.3. Inclusion criteria for baseline (2016) and last follow-up year (2021), by CPC+ versus comparison status, Track 1

	2016 CPC+		2016 comparison		2021 CPC+		2021 comparison	
	Sample size	Percentage of beneficiaries retained from one step to the next	Sample size	Percentage of beneficiaries retained from one step to the next	Sample size	Percentage of beneficiaries retained from one step to the next	Sample size	Percentage of beneficiaries retained from one step to the next
Number of Medicare FFS beneficiaries assigned to a practice	873,870		2,899,486		1,059,086		3,763,839	
Number of practices	1,373[a]		5,243[b]		1,373[a]		5,242[b]	
Beneficiary inclusion criteria								
Continuously enrolled in Medicare Parts A, B, and D during the measurement year or until death	610,785	69.9%	1,983,103	68.4%	765,081	72.2%	2,704,286	71.8%
Any opioid use	211,246	34.6%	672,126	33.9%	185,947	**24.3%**	644,840	23.8%
Appropriate use criteria								
No cancer in measurement year or in prior year	168,630	79.8%	539,753	80.3%	143,255	77.0%	497,555	77.2%
No sickle cell in measurement year	168,518	99.9%	539,354	99.9%	143,157	99.9%	497,110	99.9%
No hospice use in measurement year	165,120	98.0%	529,248	98.1%	139,495	97.4%	484,799	97.5%
Long-term opioid use and potential overuse								
Long-term opioid use								
Age 18 or older and has a 90+ day supply of opioids in a measurement year with no more than a 7-day gap between prescriptions	40,219	24.4%	129,178	24.4%	30,037	21.5%	103,103	21.3%
Potential opioid overuse								
Beneficiaries who use opioids long term and who potentially overuse opioids	7,743	19.3%	24,285	18.8%	3,505	11.7%	12,342	12.0%
Number of practices with beneficiaries who use opioids long term[c]	1,355		5,157		1,352		5,131	

Source: Analysis of Medicare claims data for 2016 and 2021.

Notes: All the counts and corresponding percentages are raw counts, unadjusted and unweighted.

[a] We excluded from the analysis practices that withdrew from CPC+ in the first three months because they were unlikely to have made much progress implementing CPC+ during that time; there were 17 such practices in the two tracks combined.

[b] In 2021, there was one fewer comparison practice than in 2016 because that practice closed.

[c] Because potential opioid overuse is assessed among beneficiaries who use opioids long term, the analysis sample excludes CPC+ and comparison practices without assigned beneficiaries who use opioids long term.

CPC+ = Comprehensive Primary Care Plus; FFS = fee-for-service.

Table 5.G.4. Inclusion criteria for baseline (2016) and last follow-up year (2021), by CPC+ versus comparison status, Track 2

	2016 CPC+		2016 comparison		2021 CPC+		2021 comparison	
	Sample size	Percentage of beneficiaries retained from one step to the next	Sample size	Percentage of beneficiaries retained from one step to the next	Sample size	Percentage of beneficiaries retained from one step to the next	Sample size	Percentage of beneficiaries retained from one step to the next
Number of Medicare FFS beneficiaries assigned to a practice	1,066,826		2,461,805		1,307,703		3,174,445	
Number of practices	1,515[a]		3,783		1,515[a]		3,783	
Beneficiary inclusion criteria								
Continuously enrolled in Medicare Parts A, B, and D during the measurement year or until death	739,187	69.3%	1,687,030	68.5%	936,733	71.6%	2,293,814	72.3%
Any opioid use	**255,288**	**34.5%**	**566,669**	**33.6%**	**225,279**	**24.0%**	**544,094**	**23.7%**
Appropriate use criteria								
No cancer in measurement year or in prior year	204,168	80.0%	453,160	80.0%	173,991	77.2%	418,474	76.9%
No sickle cell in measurement year	204,036	99.9%	452,818	99.9%	173,842	99.9%	418,112	99.9%
No hospice use in measurement year	200,008	98.0%	444,062	98.1%	169,198	97.3%	407,646	97.5%
Long-term opioid use and potential overuse								
Long-term opioid use								
Age 18 or older and has a 90+ day supply of opioids in a measurement year with no more than a 7-day gap between prescriptions	48,744	24.4%	105,422	23.7%	36,570	21.6%	84,267	20.7%
Potential opioid overuse								
Beneficiaries who use opioids long term and who potentially overuse opioids	9,530	19.6%	20,121	19.1%	4,334	11.9%	10,126	12.0%
Number of practices with beneficiaries who use opioids long term[b]	1,500		3,733		1,495		3,696	

Source: Analysis of Medicare claims data for 2016 and 2021.

Notes: All the counts and corresponding percentages are raw counts, unadjusted and unweighted.

[a] We excluded from the analysis practices that withdrew from CPC+ in the first three months because they were unlikely to have made much progress implementing CPC+ during that time; there were 17 such practices in the two tracks combined.

[b] Because potential opioid overuse is assessed among beneficiaries who use opioids long term, the analysis sample excludes CPC+ and comparison practices without assigned beneficiaries who use opioids long term.

CPC+ = Comprehensive Primary Care Plus; FFS = fee-for-service.

A.2. Baseline equivalence

After matching, for both CPC+ tracks, CPC+ and comparison beneficiaries with long-term opioid use (those eligible for the overuse measure) had similar baseline characteristics, including similar rates of potential opioid overuse, demographics, chronic conditions, Medicare expenditures, hospitalizations, and emergency department (ED) use. CPC+ and comparison practices were also similar in size, health system ownership status, and experience with EHRs and prior primary care transformation experience; counties in which they were located were also similar on characteristics such as median income, rural/urban location, and percentage of the population in poverty.

Nearly all standardized differences in baseline practice and beneficiary characteristics were less than 0.10. Standardized differences were larger in subgroups but still within 0.25 (data available upon request). Beneficiary, practice, and market characteristics were similarly well balanced between CPC+ and comparison practices in the long-term use analysis sample. Detailed balance information is in Table 5.G.5 for the potential opioid use analysis. Highlighted cells are for standardized differences larger than 0.10. We also show detailed balance information for the long-term opioid sample, see Table 5.G.6.

Table 5.G.5. Baseline characteristics (2016) for CPC+ and comparison groups in the potential opioid overuse analysis sample, by track[a]

	Track 1			Track 2		
	Mean among CPC+ practices (N = 1,356)	Weighted mean among comparison practices (N = 5,157)	Standardized differences	Mean among CPC+ practices (N = 1,502)	Weighted mean among comparison practices (N = 3,746)	Standardized differences
Potential opioid overuse	**19.2**	**18.3**	**0.02**	**19.5**	**19.2**	**0.01**
Beneficiary characteristics						
Demographics						
Age						
18–65	45.9	45.7	0.00	45.4	45.9	-0.01
65–74	31.9	31.2	0.01	32.1	31.0	0.02
75–84	15.5	15.8	-0.01	15.4	15.7	-0.01
85+	6.8	7.4	-0.02	7.1	7.5	-0.02
Race						
White	86.4	87.7	-0.04	86.8	88.1	-0.04
Black	9.4	8.2	0.04	9.1	8.2	0.03
All other/unknown	4.3	4.1	0.01	4.1	3.7	0.02
Male	35.1	34.9	0.00	34.4	34.8	-0.01
Eligibility for Medicare and Medicaid						
Original reason for Medicare eligibility						
Disability	59.0	58.7	0.01	58.3	58.5	0.00
Age	40.1	40.4	-0.01	40.8	40.5	0.01
ESRD	0.9	0.9	0.00	0.9	1.0	-0.02
Dual eligibility	40.5	43.8	-0.07	39.8	43.3	-0.07
Presence of chronic conditions[b]						
Chronic obstructive pulmonary disease	24.6	24.5	0.00	23.4	23.5	0.00
Vascular disease, with or without complications	18.6	18.7	0.00	18.0	19.1	-0.03
Diabetes with chronic complications	17.7	17.8	0.00	18.2	17.6	0.01
Rheumatoid arthritis and inflammatory connective tissue disease or disorders of immunity	17.5	16.2	0.03	17.1	16.0	0.03
Schizophrenia, major depressive, bipolar, or paranoid disorders	17.1	18.5	-0.04	18.2	19.4	-0.03
Congestive heart failure	14.7	14.7	0.00	14.5	14.6	0.00
Diabetes without complication	13.7	13.8	0.00	12.7	13.3	-0.02
Specified heart arrhythmias	12.4	12.4	0.00	12.4	12.9	-0.01
Morbid obesity	10.9	10.8	0.00	10.8	10.9	0.00
Drug/alcohol psychosis or dependence	9.9	9.9	0.00	10.4	10.5	0.00
Alzheimer's disease or dementia	7.3	7.7	-0.01	7.7	7.6	0.00

Table 5.G.5. ***(continued)***

	Track 1			Track 2		
	Mean among CPC+ practices (N = 1,356)	Weighted mean among comparison practices (N = 5,157)	Standardized differences	Mean among CPC+ practices (N = 1,502)	Weighted mean among comparison practices (N = 3,746)	Standardized differences
Risk score[c]						
Mean HCC score	1.6	1.6	-0.01	1.6	1.6	-0.03
Beneficiaries assigned a new enrollee HCC score (i.e., HCC score was calculated based on demographic characteristics only)	8.8	8.2	0.02	9.1	8.4	0.03
High-risk beneficiary – 75th percentile	40.0	40.8	-0.02	39.8	41.1	-0.03
High-risk beneficiary – 90th percentile	21.9	22.6	-0.02	21.9	22.9	-0.03
Characteristics of the beneficiary's assigned practice[d]						
Prior transformation						
Experience in selected practice transformation activities[e]	53.0	53.0	0.00	81.0	76.5	***0.11***
Participant in SSP ACO as of January 1 of the first intervention year	49.6	47.8	0.00	40.3	40.2	0.00
Meaningful EHR use[f]						
Never attested	12.3	11.7	0.02	5.8	5.7	0.00
Attested since 2011 or 2012	73.8	74.4	-0.01	84.4	84.9	-0.02
Attested since 2013 or later	13.9	13.9	0.00	9.8	9.3	0.02
Size						
Number of primary care practitioners[g]	6.6	6.6	0.00	9.1	8.9	0.01
One to two	20.8	23.4	-0.06	13.8	15.2	-0.04
Three to five	33.9	36.3	-0.05	33.3	34.9	-0.03
Six or more	45.3	40.4	0.10	52.9	49.9	0.06
Practice size category[g]						
Small (1 to 2 practitioners)	20.2	22.5	-0.06	13.3	14.5	-0.04
Medium (3 to 24 practitioners)	74.4	72.8	0.04	77.2	77.8	-0.01
Large (25 or more practitioners)	5.4	4.7	0.03	9.5	7.7	0.06
Number of Medicare beneficiaries assigned in the baseline year	1100	1051	0.05	1332	1192	***0.11***
Ownership[g]						
Hospital ownership or health system management or ownership	56.7	55.3	0.03	61.5	60.1	0.03
Hospital-owned	30.4	30.7	-0.01	29.7	33.0	-0.07
Multispecialty[h]						
Multispecialty practice	18.5	20.2	-0.04	25.6	24.2	0.03

Table 5.G.5. ***(continued)***

	Track 1			Track 2		
	Mean among CPC+ practices (N = 1,356)	Weighted mean among comparison practices (N = 5,157)	Standardized differences	Mean among CPC+ practices (N = 1,502)	Weighted mean among comparison practices (N = 3,746)	Standardized differences
Urbanicity of practice's county (Area Resource File)[i]						
Urban	64.9	64.8	0.00	71.0	70.5	0.01
Suburban	22.2	21.9	0.01	19.4	19.1	0.01
Rural	12.9	13.2	-0.01	9.6	10.4	-0.03
Practice county socioeconomic characteristics (Area Resource File)[j]						
Median household income ($)	53370	54497	-0.08	53815	54713	-0.07
Medicare Advantage penetration rate	29.6	28.8	0.07	31.3	30.9	0.04
Percentage of adults 25 or older with a degree from a four-year college	28.8	28.5	0.03	29.1	29.1	0.01
Percentage of population in poverty	15.2	14.6	***0.12***	15.0	14.7	0.06
Area with a shortage of (primary care) health professionals	1.7	1.6	0.01	1.4	1.6	-0.02
Hospital beds in county per 10,000 population (Area Resource File)[j]						
1st quartile (fewest beds)	22.9	24.4	-0.03	27.9	26.3	0.03
2nd quartile	23.0	23.8	-0.02	20.3	21.4	-0.03
3rd quartile	24.9	24.1	0.02	22.9	22.9	0.00
4th quartile (most beds)	29.2	27.8	0.03	28.9	29.4	-0.01
U.S. census region[k]						
Northeast	17.4	22.0	***-0.12***	18.2	23.4	***-0.14***
Midwest	45.1	36.4	***0.17***	35.1	34.0	0.02
South	20.8	25.0	-0.10	28.7	25.4	0.07
West	16.8	16.7	0.01	18.0	17.2	0.02
Other characteristics						
HRR price index (CMS's Medicare Geographic Variation data, 2015)	1.03	1.04	-0.06	1.03	1.04	-0.09
Service use and expenditures						
Service use (in the baseline year per 1,000 beneficiaries, annualized)						
Acute care hospitalizations	480	481	0.00	490	484	0.01
Outpatient ED visits	1061	1091	-0.01	1039	1047	0.00
Expenditures (per beneficiary per month, $)						
Medicare Part A and B expenditures without fees	1271	1281	-0.00	1289	1295	-0.00

Sources: Data on practice size and ownership from SK&A data; data on the number and characteristics of assigned Medicare beneficiaries from Medicare Enrollment Database and claims data; data on patient-centered medical home recognition from NCQA, TJC, AAAHC, URAC, and state-specific data sources; data on SSP ACO participation from CMS's master data management (MDM) data; data on participation in MAPCP and in CPC Classic from CMS; data on meaningful use of EHR from CMS's Medicare EHR Incentive Program data; data on HRR Price Index from CMS's Medicare Geographic Variation data; county data from the Area Resource File: 2015–2016.

Table 5.G.5. ***(continued)***

[a] All values in this table are reported as percentages of either beneficiaries or practices, depending on variable (multiplied by 100), except for HCC score, number of beneficiaries assigned, median household income, HRR price index, service use, and expenditures.

[b] Chronic conditions that were prevalent for greater than 10 percent of any of the samples (Track 1, Track 2) and Alzheimer's disease/dementia are reported in this table.

[c] The HCC score in the baseline year is based on beneficiaries' diagnoses in 2015. HCC scores are a measure of risk for subsequent expenditures. CMS calculates them such that the average for the Medicare FFS population nationally is 1.0. A patient with a risk score of 1.30 is predicted to have expenditures that would be approximately 30 percent above the average, whereas a patient with a risk score of 0.70 is expected to have expenditures that would be approximately 30 percent below the average.

[d] Practice is defined as a physical location or practice site.

[e] We define prior transformation experience as CPC Classic or MAPCP participation, or NCQA, TJC, AAAHC, URAC, or state medical-home recognition status (whether practice is in a medical home). Data from 2016 on patient-centered medical home recognition from NCQA, TJC, AAAHC, URAC, and state-specific data sources. Data from 2016 on participation in MAPCP and in CPC Classic from CMS.

[f] Practice with at least one practitioner who attested to meaningful use of EHR; year of first attestation of meaningful use of EHR.

[g] Data on practice size and ownership from 2016 SK&A data.

[h] We define multispecialty as having at least one practitioner, according to SK&A, with a specialty other than general practice, internal medicine, family medicine, or geriatrics.

[i] The urbanicity of a practice's county (rural, urban, suburban) is derived from the 2013 (latest year available) rural-urban continuum codes (https://www.ers.usda.gov/data-products/rural-urban-continuum-codes/documentation/) available in the ARF.

[j] Due to lags in the ARF data, the specific year of each geographic characteristic may differ depending on the most recent year of data available. For determining whether a practice was located in a health professionals shortage area, we used data from years 2015 and 2016. For median household income, percentage of population in poverty, and Medicare Advantage penetration rate in the practice's county, we used data from 2014. For hospital beds in the practice's county, we used data from 2013 and determined county population (for creating the per 10,000 population measure of hospital beds) using 2014 data. For percentage of adults 25 or older with a degree from a four-year college, we used data from years 2010–2014.

[k] For ease of presentation, we show balance on the four Census Bureau-designated regions (based on the state of the practice) in this table. However, for inclusion in the propensity score matching model, we identified comparison market areas for each CPC+ region (or groups of regions) based on geographic proximity, the primary care landscape, and number of available potential comparison practices. Details on the selection of external regions are available in Supplement 1.A.

AAAHC = Accreditation Association for Ambulatory Health Care; ACO = Accountable Care Organization; ARF = Area Resource File; ED = emergency department; EHR = electronic health record; ESRD = end-stage renal disease; FFS = fee-for-service; HCC = Hierarchical Condition Category; HHA = home health agency; HRR = hospital referral region; MAPCP = Multi-Payer Advanced Primary Care Practice; NCQA = National Committee for Quality Assurance; QPP = Quality Payment Program; SNF = skilled nursing facility; SSP = Medicare Shared Savings Program; TJC = The Joint Commission; URAC = Utilization Review Accreditation Commission.

Table 5.G.6. Baseline characteristics (2016) for CPC+ and comparison groups in the long-term opioid use analysis sample, by track[a]

	Track 1			Track 2		
	Mean among CPC+ practices (N = 1,373)	Weighted mean among comparison practices (N = 5,243)	Standardized differences	Mean among CPC+ practices (N = 1,515)	Weighted mean among comparison practices (N = 3,782)	Standardized differences
Long-term opioid use	**8.8**	**8.6**	**0.01**	**8.9**	**8.5**	**0.01**
Beneficiary characteristics						
Demographics						
Age						
18–65	16.3	17.1	-0.02	16.2	17.6	-0.04
65–74	49.0	47.6	0.03	49.3	47.6	0.03
75–84	24.4	24.6	0.00	24.2	24.3	0.00
85 +	10.3	10.6	-0.01	10.3	10.5	-0.01
Race						
White	88.3	88.0	0.01	87.4	87.4	0.00
Black	5.7	5.7	0.00	6.3	6.2	0.00
All other/unknown	6.1	6.2	-0.01	6.3	6.3	0.00
Male	38.6	38.6	0.00	39.0	38.8	0.00
Eligibility for Medicare and Medicaid						
Original reason for Medicare eligibility						
Disability	22.7	23.4	-0.02	22.5	23.8	-0.03
Age	76.8	76.0	0.02	77.0	75.7	0.03
ESRD	0.5	0.5	-0.01	0.5	0.6	-0.01
Dual eligibility	18.1	21.7	-0.09	17.9	22.0	***-0.11***
Presence of chronic conditions[b]						
Chronic obstructive pulmonary disease	11.9	11.7	0.01	11.3	11.3	0.00
Vascular disease, with or without complications	14.6	14.7	0.00	14.5	14.5	0.00
Diabetes with chronic complications	12.4	12.5	0.00	12.8	12.7	0.00
Rheumatoid arthritis and inflammatory connective tissue disease or disorders of immunity	7.3	6.9	0.01	7.0	6.9	0.00
Schizophrenia, major depressive, bipolar, or paranoid disorders	8.5	9.5	-0.04	9.0	9.8	-0.03
Congestive heart failure	9.9	9.9	0.00	9.8	9.9	0.00
Diabetes without complication	13.1	13.0	0.00	12.0	12.5	-0.02
Specified heart arrhythmias	12.7	12.6	0.00	12.6	12.7	0.00
Morbid obesity	5.0	4.8	0.01	4.9	5.0	0.00
Drug/alcohol psychosis or dependence	2.3	2.4	-0.01	2.4	2.5	-0.01
Alzheimer's disease or dementia	6.6	6.8	-0.01	6.6	6.8	-0.01

Table 5.G.6. ***(continued)***

	Track 1			Track 2		
	Mean among CPC+ practices (N = 1,373)	Weighted mean among comparison practices (N = 5,243)	Standardized differences	Mean among CPC+ practices (N = 1,515)	Weighted mean among comparison practices (N = 3,782)	Standardized differences
Risk score[c]						
Mean HCC score	1.1	1.1	-0.01	1.1	1.1	-0.02
Beneficiaries assigned a new enrollee HCC score (i.e., HCC score was calculated based on demographic characteristics only)	6.3	5.4	0.03	7.2	5.6	0.06
High-risk beneficiary – 75th percentile	22.0	22.4	-0.01	21.9	22.6	-0.02
High-risk beneficiary – 90th percentile	12.7	13.2	-0.01	12.7	13.2	-0.01
Characteristics of the beneficiary's assigned practice[d]						
Prior transformation						
Experience in selected practice transformation activities[e]	53.0	52.3	0.01	80.9	75.1	***0.15***
Participant in SSP ACO as of January 1 of the first intervention year	51.8	52.5	-0.01	44.6	44.0	0.01
Meaningful EHR use[f]						
Never attested	8.1	8.8	-0.03	3.6	3.9	-0.02
Attested since 2011 or 2012	78.7	78.2	0.01	88.0	87.7	0.01
Attested since 2013 or later	13.2	13.0	0.01	8.4	8.3	0.00
Size						
Number of primary care practitioners[g]	6.6	6.9	-0.06	9.4	9.4	0.00
One to two	21.7	22.1	-0.01	13.2	14.0	-0.03
Three to five	32.8	34.4	-0.03	32.2	33.2	-0.02
Six or more	45.5	43.5	0.04	54.6	52.8	0.04
Practice size category[g]						
Small (1 to 2 practitioners)	21.2	21.5	-0.01	12.7	13.3	-0.02
Medium (3 to 24 practitioners)	73.9	74.0	0.00	77.4	78.2	-0.02
Large (25 or more practitioners)	5.0	4.5	0.02	10.0	8.5	0.05
Number of Medicare beneficiaries assigned in the baseline year	1,176	1,125	0.05	1,358	1,296	0.05
Ownership[g]						
Hospital ownership or health system management or ownership	54.8	55.2	-0.01	58.5	59.8	-0.03
Hospital-owned	27.9	28.6	-0.02	29.0	31.1	-0.05
Multispecialty[h]						
Multispecialty practice	19.5	19.8	-0.01	25.6	25.6	0.00

Table 5.G.6. ***(continued)***

	Track 1			Track 2		
	Mean among CPC+ practices (N = 1,373)	Weighted mean among comparison practices (N = 5,243)	Standardized differences	Mean among CPC+ practices (N = 1,515)	Weighted mean among comparison practices (N = 3,782)	Standardized differences
Urbanicity of practice's county (Area Resource File)[i]						
Urban	71.3	70.7	0.01	76.2	74.4	0.04
Suburban	18.4	18.9	-0.01	15.9	17.3	-0.04
Rural	10.3	10.4	-0.01	7.9	8.3	-0.02
Practice county socioeconomic characteristics (Area Resource File)[j]						
Median household income ($)	58,017	57,784	0.01	57,138	57,242	-0.01
Medicare Advantage penetration rate	28.3	28.4	-0.01	31.2	30.3	0.07
Percentage of adults 25 or older with a degree from a four-year college	31.5	31.0	0.04	31.2	30.9	0.02
Percentage of population in poverty	13.8	14.0	-0.02	14.2	14.2	-0.01
Area with a shortage of (primary care) health professionals	1.0	1.2	-0.03	1.3	1.4	-0.01
Hospital beds in county per 10,000 population (Area Resource File)[j]						
1st quartile (fewest beds)	21.4	21.9	-0.01	24.7	23.3	0.03
2nd quartile	28.0	25.4	0.06	24.3	23.7	0.01
3rd quartile	26.2	27.1	-0.02	24.4	26.5	-0.05
4th quartile (most beds)	24.3	25.6	-0.03	26.6	26.5	0.00
U.S. census region[k]						
Northeast	29.2	28.7	0.01	27.6	28.5	-0.02
Midwest	38.8	35.9	0.06	35.1	35.5	-0.01
South	15.0	18.7	***-0.10***	19.1	19.0	0.00
West	17.0	16.6	0.01	18.1	17.0	0.03
Other characteristics						
HRR price index (CMS's Medicare Geographic Variation data, 2015)	1.1	1.1	-0.09	1.0	1.1	-0.08
Service use and expenditures						
Service use (in the baseline year per 1,000 beneficiaries, annualized)						
Acute care hospitalizations	262.4	266.4	0.00	264.7	266.4	0.00
Outpatient ED visits	505.3	521.3	-0.01	503.0	513.9	-0.01
Expenditures (per beneficiary per month, $)						
Medicare Part A and B expenditures without fees	771.8	782.4	-0.01	764.4	777.1	-0.01

Sources: Data on practice size and ownership from SK&A data; data on the number and characteristics of assigned Medicare beneficiaries from Medicare Enrollment Database and claims data; data on patient-centered medical home recognition from NCQA, TJC, AAAHC, URAC, and state-specific data sources; data on SSP ACO participation from CMS's master data management (MDM) data; data on participation in MAPCP and in CPC Classic from CMS; data on meaningful use of EHR from CMS's Medicare EHR Incentive Program data; data on HRR Price Index from CMS's Medicare Geographic Variation data; county data from the Area Resource File: 2015–2016.

Table 5.G.6. ***(continued)***

[a] All values in this table are reported as percentages of either beneficiaries or practices, depending on variable (multiplied by 100), except for HCC score, number of beneficiaries assigned, median household income, HRR price index, and utilization and service use measures.

[b] Chronic conditions reported in this table are those that were prevalent for greater than 10 percent of any of the samples in the potential opioid overuse analysis, plus Alzheimer's disease/dementia—see Supplement Table 2.3. They are listed in the same order as in that table.

[c] The HCC score in the baseline year is based on beneficiaries' diagnoses in 2015. HCC scores are a measure of risk for subsequent expenditures. CMS calculates them such that the average for the Medicare FFS population nationally is 1.0. A patient with a risk score of 1.30 is predicted to have expenditures that would be approximately 30 percent above the average, whereas a patient with a risk score of 0.70 is expected to have expenditures that would be approximately 30 percent below the average.

[d] Practice is defined as a physical location or practice site.

[e] We define prior transformation experience as CPC Classic or MAPCP participation, or NCQA, TJC, AAAHC, URAC, or state medical-home recognition status (whether practice is in a medical home). Data from 2016 on patient-centered medical home recognition from NCQA, TJC, AAAHC, URAC, and state-specific data sources. Data from 2016 on participation in MAPCP and in CPC Classic from CMS.

[f] Practice with at least one practitioner who attested to meaningful use of EHR; year of first attestation of meaningful use of EHR.

[g] Data on practice size and ownership from 2016 SK&A data.

[h] We define multispecialty as having at least one practitioner, according to SK&A, with a specialty other than general practice, internal medicine, family medicine, or geriatrics.

[i] The urbanicity of a practice's county (rural, urban, suburban) is derived from the 2013 (latest year available) rural-urban continuum codes (https://www.ers.usda.gov/data-products/rural-urban-continuum-codes/documentation/) available in the ARF.

[j] Due to lags in the ARF data, the specific year of each geographic characteristic may differ depending on the most recent year of data available. For determining whether a practice was located in a health professionals shortage area, we used data from years 2015 and 2016. For median household income, percentage of population in poverty, and Medicare Advantage penetration rate in the practice's county, we used data from 2014. For hospital beds in the practice's county, we used data from 2013 and determined county population (for creating the per 10,000 population measure of hospital beds) using 2014 data. For percentage of adults 25 or older with a degree from a four-year college, we used data from years 2010–2014.

[k] For ease of presentation, we show balance on the four Census Bureau-designated regions (based on the state of the practice) in this table. However, for inclusion in the propensity score matching model, we identified comparison market areas for each CPC+ region (or groups of regions) based on geographic proximity, the primary care landscape, and number of available potential comparison practices. Details on the selection of external regions are available in Supplement 1.A.

AAAHC = Accreditation Association for Ambulatory Health Care; ACO = Accountable Care Organization; ARF = Area Resource File; ED = emergency department; EHR = electronic health record; ESRD = end-stage renal disease; FFS = fee-for-service; HCC = Hierarchical Condition Category; HHA = home health agency; HRR = hospital referral region; MAPCP = Multi-Payer Advanced Primary Care Practice; NCQA = National Committee for Quality Assurance; QPP = Quality Payment Program; SNF = skilled nursing facility; SSP = Medicare Shared Savings Program; TJC = The Joint Commission; URAC = Utilization Review Accreditation Commission.

A.3. Sample characteristics

About 44 percent of beneficiaries who used opioids long term were under 65 years old; a similar proportion was dually eligible for Medicare and Medicaid, and close to 60 percent were enrolled in Medicare because of a disability. Beneficiaries who used opioids long term were younger, and a larger proportion was dually eligible for Medicare and Medicaid and enrolled in Medicare because of a disability, relative to beneficiaries who did not use opioids long term. Beneficiaries using opioids long term had, on average, many more chronic conditions and much higher rates of service use relative to those who did not use opioids long term (Tables 5.G.7 and 5.G.8).

Beneficiaries with potential opioid overuse were younger, much more likely to have a disability and a diagnosis for drug/alcohol psychosis or dependence than were beneficiaries without opioid overuse. Beneficiaries with potential overuse also had 20 percent higher hospitalization rates than those who did not. These beneficiary characteristics were similar in both tracks and for CPC+ and for comparison beneficiaries (Tables 5.G.7 and 5.G.8).

We found no differences in changes in beneficiary sample characteristics between 2016 and 2021 for CPC+ versus comparison beneficiaries included in the impact analysis in either track, which gives us confidence that estimated impacts are not driven by compositional differences between baseline and follow-up (Figure 5.G.1).

Table 5.G.7. Key baseline characteristics (2016) for CPC+ beneficiaries by whether they use opioids long term or potentially overuse them, Track 1

	Not using long term (N = 570,659)	Using long term (N = 40,219)	Using long term but not at a high dosage (N = 32,479)	Potentially overusing (N = 7,743)
Beneficiary characteristics				
Demographics (percentage)				
Age				
18–65	12.1	43.5	38.6	64.0
65–74	49.1	32.9	34.7	25.5
75–84	26.8	16.4	18.4	8.0
85+	12.0	7.2	8.4	2.5
Race				
White	88.8	87.0	86.8	87.7
Black	5.3	8.8	9.0	7.7
All other/unknown	5.9	4.2	4.1	4.6
Male	40.6	34.9	33.4	41.3
Eligibility for Medicare and Medicaid (percentage)				
Original reason for Medicare eligibility				
Age	81.2	41.0	46.0	20.4
Disability	18.4	58.0	53.1	78.8
ESRD	0.5	0.9	0.9	0.8
Dual eligibility	15.2	40.7	39.5	45.8
Presence of chronic conditions[a] (percentage)				
Vascular disease, with or without complications	16.1	19.3	19.7	17.8
Chronic obstructive pulmonary disease	12.2	25.3	24.8	27.7
Diabetes with chronic complications	12.5	18.2	18.5	17.2
Congestive heart failure	10.9	15.3	15.6	13.8
Schizophrenia, major depressive, bipolar, or paranoid disorders	7.5	17.3	16.3	21.5
Rheumatoid arthritis or disorders of immunity	7.1	18.1	17.3	21.4
Drug/alcohol psychosis or dependence	1.7	10.1	8.1	18.5
Risk score[b]				
Mean HCC score	1.2	1.6	1.6	1.7
Service use and expenditures				
Service use (in the baseline year per 1,000 beneficiaries, annualized)				
Acute care hospitalizations	341	476	458	553
Outpatient ED visits	499	1,031	1,018	1,086
Expenditures (per beneficiary per month, $)				
Medicare Part A and B expenditures without fees	1,002	1,268	1,236	1,402
Key characteristics of the beneficiary's assigned practice				
Urbanicity of practice's county (Area Resource File)[c] (percentage)				
Urban	72.3	64.3	62.7	70.7
Suburban	17.9	22.5	23.3	19.0
Rural	9.7	13.2	13.9	10.3
U.S. Census region[d] (percentage)				
Northeast	31.3	17.6	16.0	24.1
Midwest	37.9	44.7	46.2	38.1
South	14.1	20.8	21.6	17.5
West	16.7	17.0	16.2	20.3

Source: Medicare Enrollment Database and claims data for 2014 through 2016.

[a] Chronic conditions that were prevalent for greater than 15 percent in any of the four samples (not using long term; using long term; using long term but not at high dosage; potentially overusing) are reported in this table.

[b] The HCC score in the baseline year is based on beneficiaries' diagnoses in 2015. HCC scores are a measure of risk for subsequent expenditures. CMS calculates them such that the average for the Medicare FFS population nationally is 1.0. A patient with a risk score of 1.30 is predicted to have expenditures that would be approximately 30 percent above the average, whereas a patient with a risk score of 0.70 is expected to have expenditures that would be approximately 30 percent below the average.

Table 5.G.7. ***(continued)***

[c] We derive the urbanicity of a practice's county (rural, urban, suburban) from the 2013 (latest year available) rural-urban continuum codes (https://www.ers.usda.gov/data-products/rural-urban-continuum-codes/documentation/) available in the Area Resource File.

[d] We show the proportion of practices located in each of the four U.S. Census regions.

CMS = Centers for Medicare & Medicaid Services; CPC+ = Comprehensive Primary Care Plus; ED = emergency department; FFS = fee-for-service; ESRD = end-stage renal disease; HCC = Hierarchical Condition Category.

Table 5.G.8. Key baseline characteristics (2016) for CPC+ beneficiaries by whether they use opioids long term or potentially overuse them, Track 2

	Not using long term (N = 690,590)	Using long term (N = 48,747)	Using long term, but not at high dosage (N = 39,216)	Potentially overusing (N = 9,531)
Beneficiary characteristics				
Demographics (percentage)				
Age				
18–65	12.1	42.8	37.7	63.6
65–74	49.4	33.3	35.2	25.8
75–84	26.6	16.4	18.4	7.9
85+	11.9	7.5	8.7	2.8
Race				
White	87.9	87.3	87.3	87.3
Black	5.9	8.6	8.7	8.0
All other/unknown	6.2	4.1	3.9	4.7
Male	41.0	34.1	32.5	41.0
Eligibility for Medicare and Medicaid (percentage)				
Original reason for Medicare eligibility				
Age	81.3	42.1	47.3	20.6
Disability	18.3	57.1	51.8	78.6
ESRD	0.5	0.9	0.9	0.8
Dual eligibility	15.1	39.6	37.8	47.2
Presence of chronic conditions[a] (percentage)				
Vascular disease, with or without complications	15.9	18.7	19.0	17.5
Chronic obstructive pulmonary disease	11.6	23.9	23.4	25.8
Diabetes with chronic complications	12.9	18.7	18.9	17.8
Congestive heart failure	10.8	15.1	15.3	14.2
Schizophrenia or major depressive, bipolar, and paranoid disorders	7.9	18.4	17.4	22.5
Rheumatoid arthritis and inflammatory connective tissue disease or disorders of immunity	6.8	17.7	17.0	20.8
Drug/alcohol psychosis or dependence	1.7	10.5	8.5	19.1
Risk score[b]				
Mean HCC score	1.2	1.6	1.5	1.7
Service use and expenditures				
Service use (in the baseline year per 1,000 beneficiaries, annualized)				
Acute care hospitalizations	341	487	469	558
Outpatient ED visits	497	1,010	976	1,146
Expenditures (per beneficiary per month, $)				
Total Part A and B expenditures without fees	991	1,288	1,258	1,411
Key characteristics of the beneficiary's assigned practice				
Urbanicity of practice's county (Area Resource File)[c] (percentage)				
Urban	77.0	70.6	69.8	73.5
Suburban	15.5	19.6	20.1	17.6
Rural	7.5	9.8	10.1	8.8
U.S. Census Region[d] (percentage)				
Northeast	29.5	18.4	16.6	25.7
Midwest	35.0	34.3	36.3	26.2
South	17.8	29.0	29.5	27.2
West	17.7	18.3	17.7	20.8

Source: Medicare Enrollment Database and claims data for 2014 through 2016.

[a] Chronic conditions that were prevalent for greater than 10 percent of any of the sample categories (not using long term; using long term; using long term, but not at high dosage; potentially overusing) and Alzheimer's disease/dementia are reported in this table.

Table 5.G.8. ***(continued)***

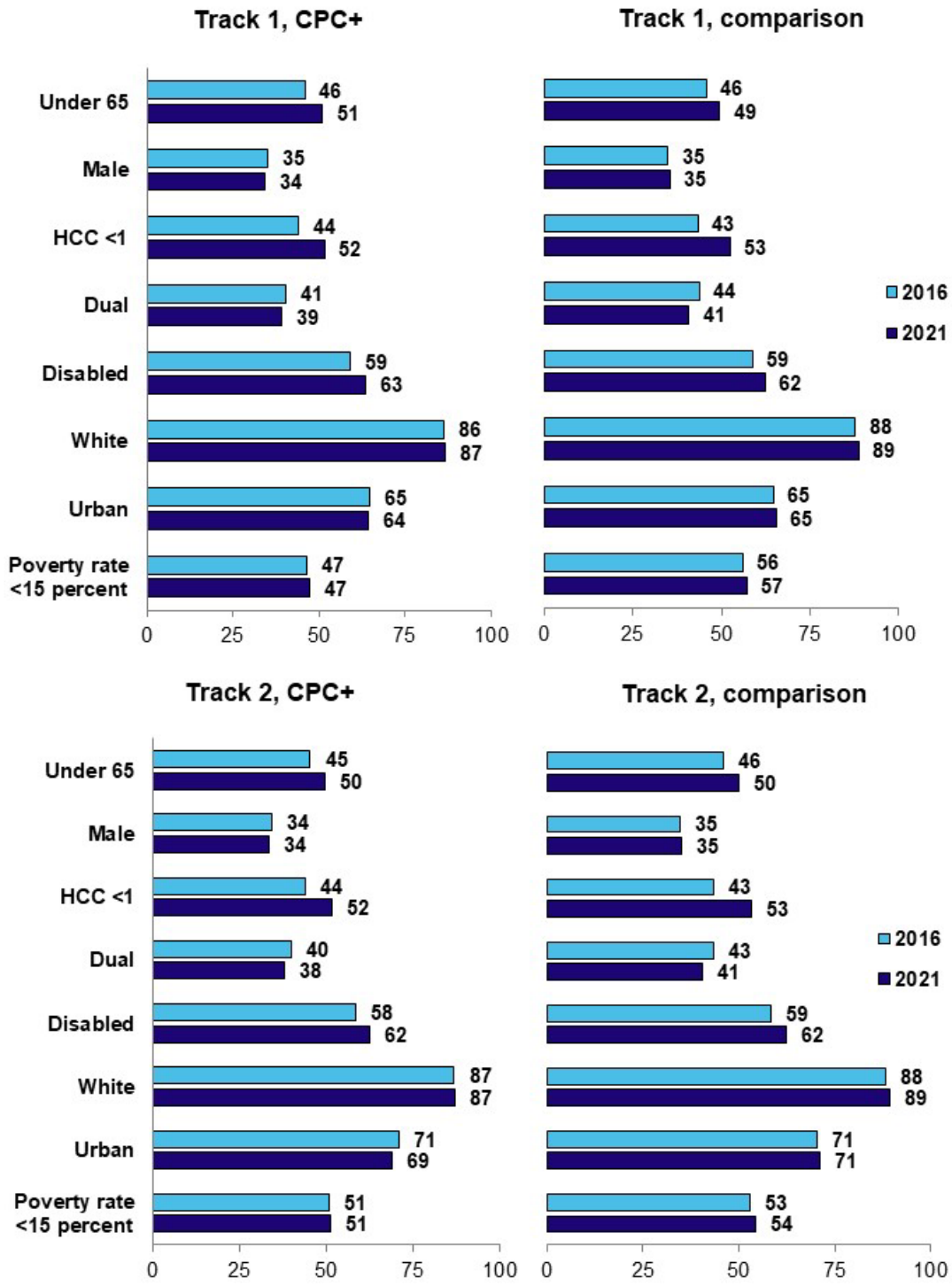
[b] The HCC score in the baseline year is based on beneficiaries' diagnoses in 2015. HCC scores are a measure of risk for subsequent expenditures. CMS calculates them such that the average for the Medicare FFS population nationally is 1.0. A patient with a risk score of 1.30 is predicted to have expenditures that would be approximately 30 percent above the average, whereas a patient with a risk score of 0.70 is expected to have expenditures that would be approximately 30 percent below the average.

[c] We derive the urbanicity of a practice's county (rural, urban, suburban) from the 2013 (latest year available) rural-urban continuum codes (https://www.ers.usda.gov/data-products/rural-urban-continuum-codes/documentation/) available in the Area Resource File.

[d] We show the proportion of practices located in each of the four U.S. Census regions.

CMS = Centers for Medicare & Medicaid Services; CPC+ = Comprehensive Primary Care Plus; ED = emergency department; ESRD = end-stage renal disease; FFS = fee-for-service; HCC = Hierarchical Condition Category.

Figure 5.G.1. Changes between 2016 and 2021 in the characteristics of beneficiaries eligible for the potential overuse analysis, by track and CPC+ versus comparison status



Source: Analysis of Medicare claims data and Medicare Enrollment Database from January 2013 through December 2020 and county data from the Area Resource File 2015–2016.

Notes: All values in this figure are reported as percentages (multiplied by 100). For poverty rate, we reported the proportion of beneficiaries who live in counties with poverty rate less than 15 percent, which is roughly the mean in 2016 among the CPC+ beneficiaries in both tracks.

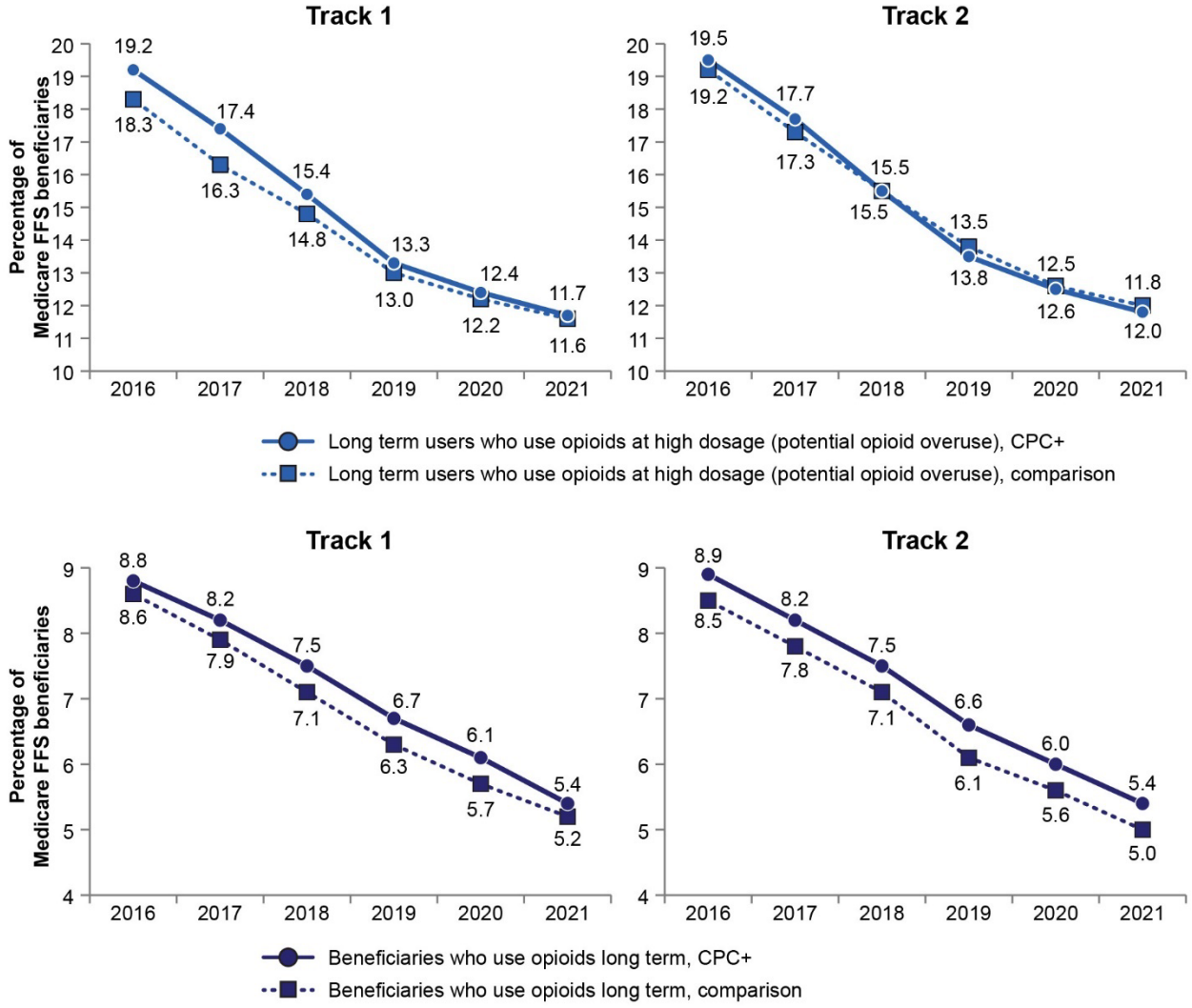
CPC+ = Comprehensive Primary Care Plus; HCC = Hierarchical Condition Category.

B. Effect of CPC+ on opioid use

B.1. Changes in regression-adjusted means

From 2016 to 2021, difference-in-differences regression-adjusted means for potential opioid overuse decreased from about 19 percent to 12 percent for CPC+ and comparison groups combined. Potential overuse started slightly higher among beneficiaries attributed to CPC+ practices in 2016 than those attributed to comparison group practices and declined to similar levels by 2021. The patterns were similar for Track 1 and Track 2 practices. Long-term opioid use also decreased over time, but to a similar extent for CPC+ and comparison group beneficiaries (Figure 5.G.2). The trends for unadjusted means were similar to those for regression-adjusted means for both outcomes and for both tracks (data available upon request).

Figure 5.G.2. Change in regression-adjusted means in long-term opioid use and potential overuse between 2016 and 2021, by track and CPC+ versus comparison status



CPC+ = Comprehensive Primary Care Plus; FFS = fee-for-service.

B.2. Estimated impacts on potential opioid overuse

Regression-adjusted difference-in-differences estimates show a larger decrease in potential opioid overuse between baseline (2016) and PY 3 through PY 5 (2019 to 2021) among beneficiaries attributed to CPC+ versus comparison practices. Between the baseline year and PY 3, the proportion of beneficiaries with potential opioid overuse decreased by 0.8 percentage points more among CPC+ Track 1 beneficiaries than among comparison group beneficiaries (p-value = 0.01) and by a similar magnitude from baseline to PY 4 (0.9 percentage points, p-value < 0.01) and from baseline to PY 5 (1.1 percentage points, p-value < 0.01). Across all five years, the estimate was favorable for CPC+ Track 1 practices (0.4 percentage points, p-value = 0.07). For Track 2, impact estimates were similar in magnitude to the Track 1 findings in PY 3, PY 4, and PY 5. The confidence intervals for the impact estimates in the two tracks overlapped to a high degree, indicating no clear evidence of differences by Track (Table 5.G.9).

Across all five program years, CPC+ reduced potential opioid overuse by 0.4 percentage points or 3.1 percent among 195,744 beneficiary-years, which translates to a reduction of 876 cases of potential opioid overuse in Track 1. For both tracks combined and across the entire intervention period, CPC+ reduced the number of cases of potential overuse by 1,817 (Table 5.G.10).

Estimated impacts on potential opioid overuse differed in significance and magnitude in some years for practices also participating in SSP relative to estimated impacts on their non-SSP counterparts; however, the confidence intervals largely overlapped (Table 5.G.9). Consistent with the overall findings, estimated impacts were the most favorable in the last two to three years of the program regardless of SSP status.

Falsification test findings showed that CPC+ and comparison practices were unlikely to have experienced differential changes in potential opioid overuse before CPC+ started (Table 5.G.11).

Across all five years of CPC+, we found no clear evidence that estimated impacts on potential opioid overuse differed among beneficiaries who have disabilities versus those who do not and dually eligible versus non–dually eligible beneficiaries (Table 5.G.12).

Table 5.G.9. Difference-in-differences estimates of the effect of CPC+ on potential opioid overuse over the five program years, by track and by SSP status

	Overall					SSP					Non-SSP				
	CPC+ mean[a]	C mean[a]	Impact estimate[b] (SE)	90 percent confidence interval	*p*-value	CPC+ mean[a]	C mean[a]	Impact estimate[b] (SE)	90 percent confidence interval	*p*-value	CPC+ mean[a]	C mean[a]	Impact estimate[b] (SE)	90 percent confidence interval	*p*-value
Track 1															
Baseline	19.2%	18.3%	n.a.	n.a.	n.a.	19.9%	18.9%	n.a.	n.a.	n.a.	18.4%	17.8%	n.a.	n.a.	n.a.
PY 1	17.4%	16.2%	0.3 (0.2)	(-0.1, 0.7)	0.26	18.2%	17.2%	0.0 (0.3)	(-0.5, 0.5)	0.92	16.5%	15.3%	0.5 (0.3)	(-0.1, 1.1)	0.15
PY 2	15.4%	15.0%	-0.5 (0.3)	(-0.9, 0.0)	0.11	16.2%	16.5%	-1.4*** (0.4)	(-2.1, -0.8)	0.00	14.7%	13.6%	0.4 (0.4)	(-0.3, 1.2)	0.32
PY 3	13.3%	13.2%	-0.8** (0.3)	(-1.3, -0.3)	0.01	14.5%	14.5%	-1.0** (0.4)	(-1.8, -0.3)	0.02	12.1%	12.0%	-0.6 (0.5)	(-1.4, 0.2)	0.20
PY 4	12.4%	12.5%	-0.9*** (0.3)	(-1.5, -0.4)	0.01	13.8%	14.0%	-1.2** (0.5)	(-2.0, -0.4)	0.01	11.1%	11.2%	-0.8 (0.5)	(-1.6, 0.1)	0.13
PY 5	11.7%	11.9%	-1.1*** (0.4)	(-1.7, -0.5)	0.00	13.0%	13.2%	-1.2** (0.5)	(-2.1, -0.3)	0.02	10.4%	10.9%	-1.1** (0.5)	(-2.0, -0.3)	0.03
PY 1 through PY 4	14.2%	13.8%	-0.4* (0.2)	(-0.9, 0.0)	0.07	15.3%	15.0%	-0.8** (0.3)	(-1.3, -0.2)	0.02	13.1%	12.6%	-0.1 (0.4)	(-0.8, 0.5)	0.71
Unweighted sample sizes[c]															
Number of beneficiaries	96,093	314,248				46,908	172,431				49,292	142,465			
Number of beneficiary-years	239,749	782,554				116,435	427,750				123,314	354,804			
Track 2															
Baseline	19.5%	19.2%	n.a.	n.a.	n.a.	19.4%	19.1%	n.a.	n.a.	n.a.	19.5%	19.4%	n.a.	n.a.	n.a.
PY 1	17.7%	17.4%	0.1 (0.2)	(-0.3, 0.5)	0.77	18.2%	17.4%	0.5 (0.4)	(-0.1, 1.1)	0.18	17.3%	17.3%	-0.2 (0.3)	(-0.7, 0.3)	0.55
PY 2	15.5%	15.7%	-0.4 (0.3)	(-0.9, 0.1)	0.19	16.4%	16.5%	-0.4 (0.5)	(-1.2, 0.4)	0.41	15.0%	15.2%	-0.4 (0.4)	(-1.1, 0.3)	0.31
PY 3	13.5%	14.0%	-0.7** (0.3)	(-1.3, -0.2)	0.04	14.7%	14.6%	-0.2 (0.6)	(-1.1, 0.7)	0.68	12.7%	13.6%	-1.1** (0.5)	(-1.9, -0.3)	0.02
PY 4	12.5%	13.0%	-0.7* (0.4)	(-1.3, -0.1)	0.05	13.5%	14.0%	-0.8 (0.6)	(-1.8, 0.2)	0.17	11.8%	12.4%	-0.7 (0.5)	(-1.5, 0.1)	0.15
PY 5	11.8%	12.3%	-0.8* (0.4)	(-1.4, -0.1)	0.05	12.9%	13.2%	-0.7 (0.6)	(-1.7, 0.4)	0.29	11.0%	11.5%	-0.7 (0.5)	(-1.6, 0.2)	0.19
PY 1 through PY 4	14.3%	14.5%	-0.4 (0.3)	(-0.8, 0.0)	0.13	15.3%	15.1%	-0.1 (0.4)	(-0.8, 0.5)	0.73	13.7%	14.1%	-0.5 (0.3)	(-1.1, 0.0)	0.12
Unweighted sample sizes[c]															
Number of beneficiaries	116,871	255,657				47,489	121,415				69,576	134,677			
Number of beneficiary-years	290,000	635,683				116,376	301,419				173,624	334,264			

Source: Analysis of Medicare claims data from January 2014 through December 2021.

Notes: Although this table indicates which estimates are statistically significant, when we interpret evidence, we combine evidence from the magnitude of the effect, the p-values, findings on related outcomes, subgroups, sensitivity tests, and other data sources about model implementation.

Table 5.G.9. ***(continued)***

[a] We report the actual, unadjusted averages in the baseline period, which are similar for the CPC+ and comparison groups due to matching. In the intervention periods, the comparison group mean is computed by subtracting the regression-adjusted difference between the CPC+ and comparison means in each time period from the CPC+ mean in that same time period.

[b] Each impact estimate is regression adjusted using a difference-in-differences analysis that reflects the difference between the average outcome for Medicare FFS beneficiaries assigned to CPC+ practices in the five years of CPC+ and the average outcome in the baseline year, relative to the same difference over time for Medicare FFS beneficiaries assigned to comparison practices, while controlling for beneficiary characteristics, practice fixed effects, changes in state-level PDMP characteristics and opioid funding, county-level opioid marketing intensity, and COVID-19 controls for 2020 and 2021.

[c] After accounting for weights that adjust for matching, the effective sample sizes fall but are still substantial. For the comparison group, the effective sample size is 41 to 46 percent of the actual sample size. The effective sample size for the CPC+ group is 100 percent because it is not affected by the matching weights.

*/**/*** Significantly different from zero at the 0.10/0.05/0.01 level, two-tailed test.

C = comparison; CPC+ = Comprehensive Primary Care Plus; FFS = fee-for-service; n.a. = not applicable because the difference-in-differences impact estimate cannot be calculated at baseline; PDMP = prescription drug monitoring program; PY = Program Year; SE = standard error; SSP = Medicare Shared Savings Program.

Table 5.G.10. The number of cases averted based on difference-in-differences estimates of the effect of CPC+ on potential opioid overuse, by track and year

	Impact estimate[a]	Percentage impact[b]	Total number of cases averted[c]
Track 1			
PY 1	0.3	1.6%	
PY 2	-0.5	-2.9%	
PY 3	-0.8**	-5.6%	
PY 4	-0.9***	-7.1%	
PY 5	-1.1***	-8.8%	
PY 1 through PY 5	-0.4*	-3.1%*	-876
Track 2			
PY 1	0.1	0.4%	
PY 2	-0.4	-2.5%	
PY 3	-0.7**	-5.1%	
PY 4	-0.7*	-5.5%	
PY 5	-0.8*	-6.0%	
PY 1 through PY 5	-0.4	-2.7%	-941

Source: Analysis of Medicare claims data from January 2014 through December 2021.

Notes: Although this table indicates which estimates are statistically significant, when we interpret evidence, we combine evidence from the magnitude of the effect, the p-values, findings on related outcomes, subgroups, sensitivity tests, and other data sources about model implementation.

[a] Each impact estimate is regression adjusted using a difference-in-differences analysis that reflects the difference between the average outcome for Medicare FFS beneficiaries assigned to CPC+ practices in the five years of CPC+ and the average outcome in the baseline year, relative to the same difference over time for Medicare FFS beneficiaries assigned to comparison practices, while controlling for beneficiary characteristics, practice fixed effects, changes in state-level PDMP characteristics and opioid funding, county-level opioid marketing intensity, and COVID-19 controls for 2020 and 2021.

[b] We calculated percentage impacts relative to what the CPC+ means would have been in each PY in the absence of the intervention—that is, the unadjusted CPC+ mean minus the impact estimate.

[c] We computed the number of averted cases based on the impact estimate and the number of beneficiary-years for beneficiaries with potential opioid overuse in any one year.

*/**/*** Significantly different from zero at the 0.10/0.05/0.01 level, two-tailed test.

CPC+ = Comprehensive Primary Care Plus; FFS = fee-for-service; PDMP = prescription drug monitoring program; PY = Program Year.

Table 5.G.11. Difference-in-differences (falsification test) results for the effect of CPC+ on potential opioid overuse in 2016, by track and by SSP status

	Overall					SSP					Non-SSP				
	CPC+ mean[a]	C mean[a]	Impact estimate[b] (SE)	90 percent confidence interval	*p*-value	CPC+ mean[a]	C mean[a]	Impact estimate[b] (SE)	90 percent confidence interval	*p*-value	CPC+ mean[a]	C mean[a]	Impact estimate[b] (SE)	90 percent confidence interval	*p*-value
Track 1															
2015	20.4%	19.4%	n.a.	n.a.	n.a.	21.1%	19.7%	n.a.	n.a.	n.a.	19.8%	19.0%	n.a.	n.a.	n.a.
2016	19.3%	18.4%	-0.1 (0.2)	(-0.5, 0.3)	0.58	20.1%	19.1%	-0.4 (0.3)	(-0.9, 0.1)	0.17	18.7%	17.9%	0.1 (0.3)	(-0.5, 0.6)	0.84
Unweighted sample sizes[c]															
Number of beneficiaries	51,626	165,994				25,689	89,344				25,937	76,650			
Number of beneficiary-years	83,046	267,282				41,224	143,680				41,822	123,602			
Track 2															
2015	20.8%	20.3%	n.a.	n.a.	n.a.	20.9%	19.9%	n.a.	n.a.	n.a.	20.8%	20.6%	n.a.	n.a.	n.a.
2016	19.7%	19.3%	-0.2 (0.2)	(-0.6, 0.2)	0.31	19.7%	19.1%	-0.3 (0.3)	(-0.9, 0.2)	0.25	19.7%	19.5%	-0.1 (0.3)	(-0.7, 0.4)	0.62
Unweighted sample sizes[c]															
Number of beneficiaries	62,346	135,328				25,067	63,606				37,279	71,722			
Number of beneficiary-years	100,761	217,752				40,420	102,168				60,341	115,584			

Source: Analysis of Medicare claims data from January 2013 through December 2016.

Notes: Although this table indicates which estimates are statistically significant, when we interpret evidence, we combine evidence from the magnitude of the effect, the p-values, findings on related outcomes, subgroups, sensitivity tests, and other data sources about model implementation.

a We report the actual, unadjusted averages in 2015. In 2016, the comparison group mean is computed by subtracting the regression adjusted difference between the CPC+ and comparison means from the CPC+ mean.

b Each impact estimate is regression adjusted using a difference-in-differences analysis that reflects the difference between the average outcome for Medicare FFS beneficiaries assigned to CPC+ practices in 2016 and the average outcome in 2015, relative to the same difference over time for Medicare FFS beneficiaries assigned to comparison practices, while controlling for beneficiary characteristics, practice fixed effects, changes in state-level opioid funding, and county-level opioid marketing intensity. We did not include changes in state-level PDMP characteristics because we did not collect those data for 2015.

c After accounting for weights that adjust for matching, the effective sample sizes fall but are still substantial. For the comparison group, the effective sample size is 39 to 48 percent of the actual sample size. The effective sample size for the CPC+ group is 100 percent of the actual sample size because it is not affected by the matching weights.

C = comparison; CPC+ = Comprehensive Primary Care Plus; FFS = fee-for-service; n.a. = not applicable because the difference-in-differences impact estimate cannot be calculated at baseline; PDMP = prescription drug monitoring program; SE = standard error; SSP = Medicare Shared Savings Program.

Table 5.G.12. Difference-in-differences estimates of the effect of CPC+ on potential opioid overuse over the five program years, by beneficiary subgroup within each track

Subgroup definition	Number (%) of CPC+ beneficiaries in subgroup at baseline	Impact estimate (standard error)	p-value for difference in impact estimates between subgroups
Track 1			
Main analysis (all Track 1 practices)	--	**-0.4* (0.2)**	--
Beneficiaries with disabilities			
Yes	25,594 (59.0%)	-0.3 (0.37)	
No	17,811 (41.0%)	-0.7 (0.27)	0.33
Beneficiaries dually eligible for Medicare and Medicaid			
Yes	17,583 (40.5%)	-0.3 (0.40)	
No	25,822 (59.5%)	-0.5 (0.30)	0.74
Track 2			
Main analysis (all Track 2 practices)	--	**-0.4 (0.3)**	--
Beneficiaries with disabilities			
Yes	32,818 (58.3%)	-0.5 (0.39)	
No	23,483 (41.7%)	-0.4 (0.28)	0.79
Beneficiaries dually eligible for Medicare and Medicaid			
Yes	22,414 (39.8%)	-0.8 (0.42)	
No	33,887 (60.2%)	-0.1 (0.30)	0.15

Source: Mathematica's analysis of Medicare claims data from January 2013 through December 2021.

Note: The *p*-values in the last column represent results from testing for statistically significant differences in impact estimates between the subgroups defined at baseline (using a t-test for subgroups with two categories). Because this test did not indicate a statistically significant or meaningful difference between any subgroups defined by the same characteristic, we did not further test whether estimates within each subgroup were statistically significant.

* Significantly different from zero at the 0.10 level, two-tailed test.

CPC+ = Comprehensive Primary Care Plus.

B.3. Estimated impacts on long-term opioid use

We found no clear evidence that CPC+ substantially affected long-term opioid use in either track, overall. The estimated impacts were small in magnitude regardless of SSP status in both tracks (Table 5.G.13).

Even though falsification test results showed a statistically significantly different change at baseline (from 2015 to 2016) for CPC+ versus comparison practices in both tracks, the estimated differences were very small in magnitude, about 0.1 percentage points (Table 5.G.14). Given large sample sizes in this analysis (about one million beneficiary-year observations), it is possible for small differences to be statistically significant. Because of a failed falsification test, the estimates for long term use are less robust. However, this does not pose a risk to our findings because the estimates of "impact" at baseline and during the intervention period are both very small. Further, the effect of CPC+ on long-term use is not our focus. The purpose of this analysis was to rule out larger reductions in long-term use in the comparison group versus the CPC+ group.

Table 5.G.13. Difference-in-differences estimates of the effect of CPC+ on long term opioid use over the five program years, by track and by SSP status

	Overall					SSP					Non-SSP				
	CPC+ mean[a]	C mean[a]	Impact estimate[b] (SE)	90 percent confidence interval	*p*-value	CPC+ mean[a]	C mean[a]	Impact estimate[b] (SE)	90 percent confidence interval	*p*-value	CPC+ mean[a]	C mean[a]	Impact estimate[b] (SE)	90 percent confidence interval	*p*-value
Track 1															
Baseline	8.8%	8.6%	n.a.	n.a.	n.a.	8.4%	7.8%	n.a.	n.a.	n.a.	9.2%	9.4%	n.a.	n.a.	n.a.
PY 1	8.2%	7.9%	0.1** (0.1)	(0.0, 0.2)	0.03	7.8%	7.3%	0.0 (0.1)	(-0.2, 0.1)	0.51	8.6%	8.5%	0.3*** (0.1)	(0.2, 0.4)	0.00
PY 2	7.5%	7.2%	0.1 (0.1)	(0.0, 0.2)	0.26	7.1%	6.6%	-0.2* (0.1)	(-0.3, 0.0)	0.08	7.9%	7.7%	0.4*** (0.1)	(0.2, 0.6)	0.00
PY 3	6.7%	6.5%	0.0 (0.1)	(-0.1, 0.1)	1.00	6.3%	6.1%	-0.3*** (0.1)	(-0.5, -0.1)	0.00	7.1%	6.9%	0.4*** (0.1)	(0.2, 0.6)	0.00
PY 4	6.1%	5.9%	-0.1 (0.1)	(-0.2, 0.1)	0.45	5.8%	5.6%	-0.4*** (0.1)	(-0.6, -0.2)	0.00	6.4%	6.3%	0.3** (0.1)	(0.0, 0.5)	0.05
PY 5	5.4%	5.4%	-0.2** (0.1)	(-0.4, -0.1)	0.03	5.1%	5.0%	-0.5*** (0.1)	(-0.7, -0.3)	0.00	5.7%	5.8%	0.1 (0.1)	(-0.1, 0.3)	0.53
PY 1 through PY 5	6.7%	6.5%	0.0 (0.1)	(-0.1, 0.1)	0.99	6.4%	6.0%	-0.2*** (0.1)	(-0.4, -0.1)	0.01	7.1%	7.0%	0.3*** (0.1)	(0.1, 0.5)	0.00
Unweighted sample sizes[c]															
Number of beneficiaries	992,417	3,392,578				512,852	1,985,920				480,933	1,414,950			
Number of beneficiary-years	3,409,652	11,593,878				1,741,448	6,767,979				1,668,204	4,825,899			
Track 2															
Baseline	8.9%	8.5%	n.a.	n.a.	n.a.	8.0%	7.7%	n.a.	n.a.	n.a.	9.5%	9.0%	n.a.	n.a.	n.a.
PY 1	8.2%	7.7%	0.1 (0.1)	(0.0, 0.2)	0.20	7.3%	7.1%	-0.1 (0.1)	(-0.2, 0.1)	0.42	8.9%	8.2%	0.2** (0.1)	(0.0, 0.3)	0.03
PY 2	7.5%	7.1%	0.0 (0.1)	(-0.1, 0.2)	0.54	6.7%	6.5%	-0.1 (0.1)	(-0.3, 0.1)	0.43	8.1%	7.5%	0.2 (0.1)	(0.0, 0.3)	0.11
PY 3	6.6%	6.3%	-0.1 (0.1)	(-0.2, 0.1)	0.48	5.9%	5.8%	-0.2 (0.1)	(-0.4, 0.0)	0.20	7.2%	6.7%	0.0 (0.1)	(-0.2, 0.2)	0.93
PY 4	6.0%	5.7%	-0.1 (0.1)	(-0.2, 0.1)	0.32	5.4%	5.3%	-0.2 (0.1)	(-0.4, 0.1)	0.23	6.5%	6.1%	-0.1 (0.1)	(-0.2, 0.1)	0.65
PY 5	5.4%	5.1%	-0.1 (0.1)	(-0.3, 0.0)	0.12	4.8%	4.8%	-0.2 (0.1)	(-0.5, 0.0)	0.12	5.8%	5.4%	-0.1 (0.1)	(-0.3, 0.1)	0.52
PY 1 through PY 4	6.7%	6.3%	0.0 (0.1)	(-0.1, 0.1)	0.75	6.0%	5.8%	-0.1 (0.1)	(-0.3, 0.0)	0.20	7.2%	6.7%	0.1 (0.1)	(-0.1, 0.2)	0.54
Unweighted sample sizes[c]															
Number of beneficiaries	1,205,132	2,861,935				542,941	1,440,270				664,340	1,427,908			
Number of beneficiary-years	4,144,495	9,821,453				1,854,357	4,941,860				2,290,138	4,879,593			

Source: Analysis of Medicare claims data from January 2014 through December 2021.

Notes: Although this table indicates which estimates are statistically significant, when we interpret evidence, we combine evidence from the magnitude of the effect, the p-values, findings on related outcomes, subgroups, sensitivity tests, and other data sources about model implementation.

Table 5.G.13. ***(continued)***

[a] We report the actual, unadjusted averages in the baseline period, which are similar for the CPC+ and comparison groups due to matching. In the intervention periods, the comparison group mean is computed by subtracting the regression-adjusted difference between the CPC+ and comparison means in each time period from the CPC+ mean in that same time period.

[b] Each impact estimate is regression adjusted using a difference-in-differences analysis that reflects the difference between the average outcome for Medicare FFS beneficiaries assigned to CPC+ practices in the five years of CPC+ and the average outcome in the baseline year, relative to the same difference over time for Medicare FFS beneficiaries assigned to comparison practices, while controlling for beneficiary characteristics, practice fixed effects, changes in state-level PDMP characteristics and opioid funding, county-level opioid marketing intensity, and COVID-19 controls for 2020 and 2021.

[c] After accounting for weights that adjust for matching, the effective sample sizes fall but are still substantial. For the comparison group, the effective sample size is 42 to 47 percent of the actual sample size. The effective sample size for the CPC+ group is 100 percent because it is not affected by the matching weights.

*/**/*** Significantly different from zero at the 0.10/0.05/0.01 level, two-tailed test.

C = comparison; CPC+ = Comprehensive Primary Care Plus; FFS = fee for service; n.a. = not applicable because the difference-in-differences impact estimate cannot be calculated at baseline; PDMP = prescription drug monitoring program; PY = Program Year; SE = standard error; SSP = Medicare Shared Savings Program.

Table 5.G.14. Difference-in-differences (falsification test) results for the effect of CPC+ on long-term opioid use in 2016, by track and by SSP status

	Overall					SSP					Non-SSP				
	CPC+ mean[a]	C mean[a]	Impact estimate[b] (SE)	90 percent confidence interval	*p*-value	CPC+ meana	C meana	Impact estimateb (SE)	90 percent confidence interval	p-value	CPC+ meana	C meana	Impact estimateb (SE)	90 percent confidence interval	p-value
Track 1															
2015	9.0%	8.8%	n.a.	n.a.	n.a.	8.6%	8.1%	n.a.	n.a.	n.a.	9.5%	9.5%	n.a.	n.a.	n.a.
2016	8.6%	8.4%	-0.1** (0.0)	(-0.2, 0.0)	0.05	8.2%	7.7%	0.0 (0.1)	(-0.1, 0.1)	0.97	9.0%	9.2%	-0.2*** (0.1)	(-0.3, 0.0)	0.01
Unweighted sample sizes[c]															
Number of beneficiaries	516,828	1,709,055				268,730	989,781				248,098	719,274			
Number of beneficiary-years	943,105	3,110,266				490,591	1,798,455				452,514	1,311,811			
Track 2															
2015	9.1%	8.6%	n.a.	n.a.	n.a.	8.2%	8.0%	n.a.	n.a.	n.a.	9.8%	9.1%	n.a.	n.a.	n.a.
2016	8.7%	8.3%	-0.1** (0.0)	(-0.2, 0.0)	0.02	7.9%	7.5%	0.1 (0.1)	(0.0, 0.2)	0.26	9.4%	8.9%	-0.2*** (0.1)	(-0.3, -0.1)	0.00
Unweighted sample sizes[c]															
Number of beneficiaries	621,544	1,449,037				276,480	725,728				345,064	723,309			
Number of beneficiary-years	1,133,087	2,639,652				502,705	1,319,972				630,382	1,319,680			

Source: Analysis of Medicare claims data from January 2013 through December 2016.

Notes: Although this table indicates which estimates are statistically significant, when we interpret evidence, we combine evidence from the magnitude of the effect, the *p*-values, findings on related outcomes, subgroups, sensitivity tests, and other data sources about model implementation.

[a] We report the actual, unadjusted averages in 2015. In 2016, the comparison group mean is computed by subtracting the regression adjusted difference between the CPC+ and comparison means from the CPC+ mean.

[b] Each impact estimate is regression adjusted using a difference-in-differences analysis that reflects the difference between the average outcome for Medicare FFS beneficiaries assigned to CPC+ practices in 2016 and the average outcome in 2015, relative to the same difference over time for Medicare FFS beneficiaries assigned to comparison practices, while controlling for beneficiary characteristics, practice fixed effects, changes in state-level opioid funding, and county-level opioid marketing intensity. We did not include changes in state-level PDMP characteristics because we did not collect those data for 2015.

[c] After accounting for weights that adjust for matching, the effective sample sizes fall but are still substantial.

/**/*** Significantly different from zero at the 0.05/0.01 level, two-tailed test.

C = comparison; CPC+ = Comprehensive Primary Care Plus; FFS = fee-for-service; n.a. = not applicable because the difference-in-differences impact estimate cannot be calculated at baseline; PDMP = prescription drug monitoring program; SE = standard error; SSP = Medicare Shared Savings Program.

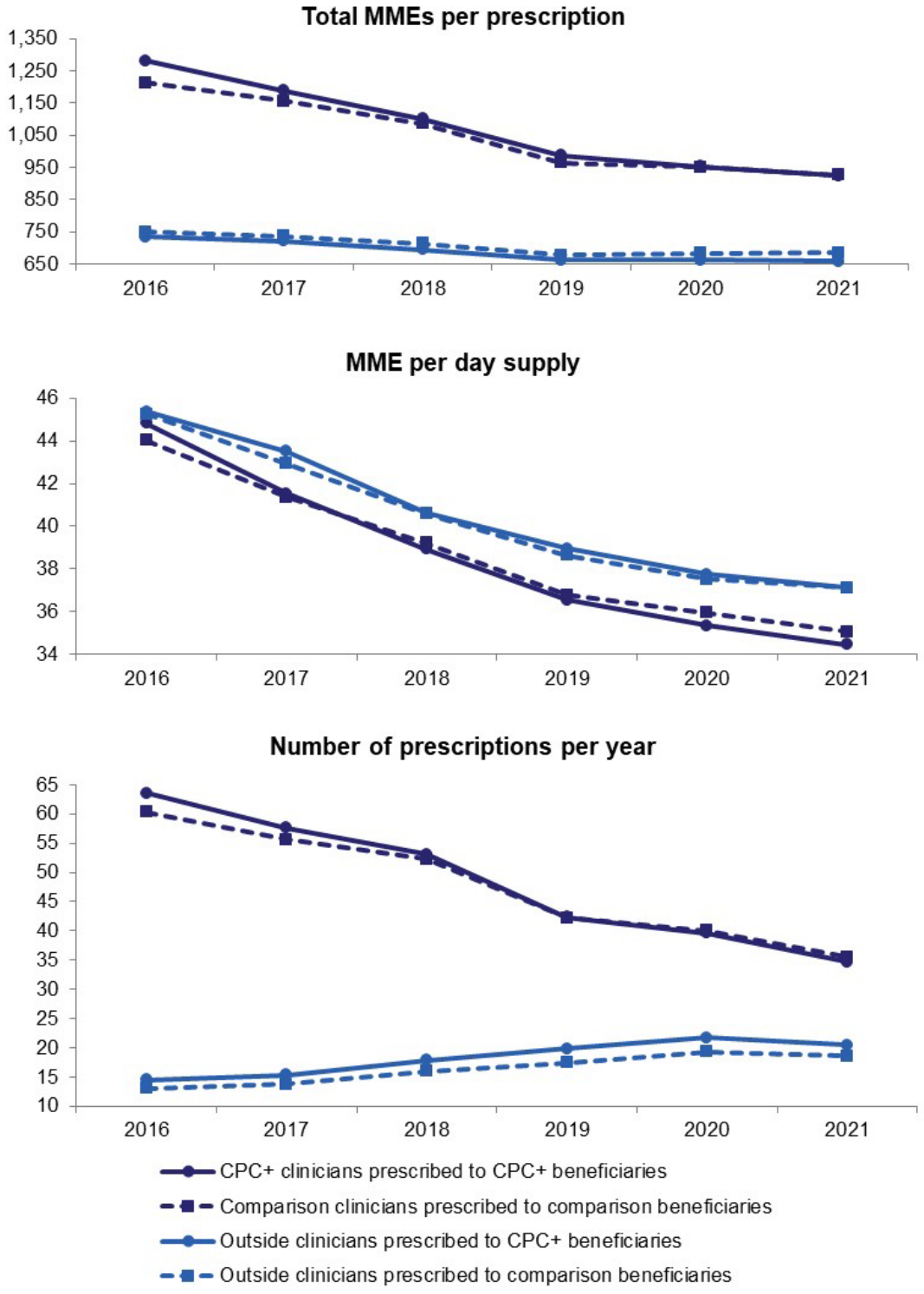
C. Effect of CPC+ on prescribing to beneficiaries who use opioids long term

C.1. Unadjusted means

This difference was driven by CPC+ and comparison clinicians prescribing nearly twice as many days of supply than did outside prescribers: about 30 days versus 16 days per prescription, respectively (Figure 5.G.3, panel 2). Further, CPC+ and comparison clinicians wrote four times as many prescriptions per year: in 2016, more than 60 prescriptions compared to about 14 for outside prescribers, on average (Figure 5.G.3, panel 3).

We ruled out that changes in outside prescribing drove the favorable impacts on opioid overuse for beneficiaries attributed to CPC+ practices. Between baseline and 2021 (PY 5), CPC+ and comparison clinicians reduced total dosage per prescription by about a quarter, whereas outside prescribers reduced it by only about 10 percent (Figure 5.G.3, panel 1). Further, although CPC+ and comparison clinicians reduced the number of prescriptions from baseline to 2021 (PY) 5 by about 7 percent, outside prescribers increased them by 10 percent (Figure 5.G.3, panel 3).

Figure 5.G.3. Prescribing by CPC+ and comparison clinicians and outside prescribers to beneficiaries with long-term opioid use, by year



Source: Mathematica's analysis of Medicare claims data from January 2013 through December 2021.
CPC+ = Comprehensive Primary Care Plus; MMEs = morphine milligram equivalents.

C.2. Estimated impacts

Based on regression-adjusted difference-in-differences results, over all five intervention years combined, CPC+ prescribers reduced total dosage and duration per prescription more than comparison prescribers. Over the entire intervention period, CPC+ Track 1 clinicians reduced opioid dosages by 37.4 MMEs per prescription (*p*-value = 0.10), or 3.5 percent more than did comparison clinicians. CPC+ clinicians reduced the duration of prescriptions by 0.5 days' supply (*p*-value = 0.02), or 1.6 percent more than did comparison clinicians (Tables 5.G.15 and 5.G.17). Estimated effects were similar in Track 2 (Tables 5.G.16 and 5.G.17).

Even though the effects on total dosage per prescription seem small, they have a much larger influence on beneficiaries' opioid use because of the large number of prescribers and prescriptions written. For example, in Track 1, more than 7,000 prescribers wrote an average of 50 prescriptions per year. In addition, CPC+ prescribers reduced the number of prescriptions per year more than comparison prescribers in Track 2 (Table 5.G.16). We found no clear evidence that Track 1 CPC+ clinicians reduced the number of prescriptions relative to the comparison clinicians (Table 5.G.15).

There were some differences in estimated impacts on prescribing outcomes by SSP status. For total dosage per prescription in both tracks and the number of days' supply in Track 2, estimated impacts were much more favorable among practices also participating in SSP. For other outcomes, the confidence intervals largely overlapped regardless of SSP status (Tables 5.G.15 and 5.G.16). Falsification tests showed no differential changes between 2015 and 2016 for prescribing outcomes (Table 5.G.18 and Table 5.G.19).

Table 5.G.15. Difference-in-differences estimates of the effect of CPC+ on opioid prescribing over the five program years among beneficiaries who use opioids long term by SSP status, Track 1

	Overall					SSP					Non-SSP				
	CPC+ mean[a]	C mean[a]	Impact estimate[b] (SE)	90 percent confidence interval	*p*-value	CPC+ mean[a]	C mean[a]	Impact estimate[b] (SE)	90 percent confidence interval	*p*-value	CPC+ mean[a]	C mean[a]	Impact estimate[b] (SE)	90 percent confidence interval	*p*-value
Total MMEs per prescription															
Baseline	1,283	1,203	n.a.	n.a.	n.a.	1,335	1,229	n.a.	n.a.	n.a.	1,229	1,175	n.a.	n.a.	n.a.
PY 1	1,190	1,141	-30.8 (22.0)	(-67.0, 5.4)	0.16	1,241	1,178	-44.1 (32.9)	(-98.3, 10.0)	0.18	1,138	1,100	-16.0 (29.6)	(-64.7, 32.7)	0.59
PY 2	1,109	1,050	-21.3 (25.7)	(-63.6, 21.1)	0.41	1,155	1,091	-42.6 (39.9)	(-108.4, 23.1)	0.29	1,063	1,015	-6.7 (34.0)	(-62.7, 49.3)	0.84
PY 3	968	926	-38.4 (27.4)	(-83.5, 6.7)	0.16	998	961	-68.9 (42.1)	(-138.1, 0.3)	0.10	937	904	-21.2 (36.6)	(-81.5, 39.0)	0.56
PY 4	954	925	-51.1* (29.3)	(-99.4, -2.9)	0.08	1,007	970	-69.6 (44.4)	(-142.6, 3.5)	0.12	902	903	-55.3 (40.3)	(-121.6, 11.0)	0.17
PY 5	915	893	-58.0** (29.5)	(-106.5, -9.6)	0.05	954	932	-84.6* (45.3)	(-159.2, -10.1)	0.06	876	867	-45.4 (38.9)	(-109.5, 18.6)	0.24
PY 1 through 5	1,026	983	-37.4* (22.5)	(-74.5, -0.4)	0.10	1,069	1,020	-57.1* (34.2)	(-113.3, -0.8)	0.10	982	954	-25.7 (30.0)	(-75.1, 23.7)	0.39
Number of days' supply per prescription															
Baseline	30.2	29.0	n.a.	n.a.	n.a.	30.6	29.1	n.a.	n.a.	n.a.	29.8	28.8	n.a.	n.a.	n.a.
PY 1	30.3	29.3	-0.3 (0.2)	(-0.6, 0.1)	0.29	30.7	29.2	0.0 (0.3)	(-0.6, 0.5)	0.93	29.9	29.4	-0.5 (0.3)	(-1.1, 0.0)	0.13
PY 2	29.7	28.9	-0.5* (0.3)	(-0.9, 0.0)	0.09	30.2	29.2	-0.5 (0.4)	(-1.2, 0.1)	0.20	29.2	28.6	-0.4 (0.4)	(-1.0, 0.2)	0.26
PY 3	27.7	27.0	-0.5* (0.3)	(-1.0, -0.1)	0.07	28.0	27.2	-0.7 (0.4)	(-1.3, 0.0)	0.11	27.5	26.8	-0.4 (0.4)	(-1.0, 0.2)	0.31
PY 4	27.9	27.3	-0.7** (0.3)	(-1.1, -0.2)	0.02	28.1	27.4	-0.8** (0.4)	(-1.5, -0.2)	0.04	27.6	26.9	-0.3 (0.4)	(-1.0, 0.3)	0.41
PY 5	27.7	27.1	-0.7** (0.3)	(-1.2, -0.2)	0.02	28.0	27.1	-0.6 (0.5)	(-1.4, 0.2)	0.20	27.4	27.2	-0.8* (0.4)	(-1.5, -0.1)	0.06
PY 1 through 5	28.6	27.8	-0.5** (0.2)	(-0.8, -0.1)	0.02	29.0	27.9	-0.4 (0.3)	(-0.9, 0.1)	0.16	28.3	27.8	-0.5 (0.3)	(-1.0, 0.0)	0.10

Table 5.G.15. ***(continued)***

	Overall					SSP					Non-SSP				
	CPC+ mean[a]	C mean[a]	Impact estimate[b] (SE)	90 percent confidence interval	*p*-value	CPC+ mean[a]	C mean[a]	Impact estimate[b] (SE)	90 percent confidence interval	*p*-value	CPC+ mean[a]	C mean[a]	Impact estimate[b] (SE)	90 percent confidence interval	*p*-value
Number of prescriptions															
Baseline	69	63	n.a.	n.a.	n.a.	66	58	n.a.	n.a.	n.a.	72	68	n.a.	n.a.	n.a.
PY 1	64	57	0.7 (1.1)	(-1.1, 2.5)	0.53	62	54	0.0 (1.4)	(-2.3, 2.3)	0.98	65	61	1.2 (1.7)	(-1.6, 4.1)	0.48
PY 2	58	51	1.4 (1.5)	(-1.0, 3.8)	0.35	56	49	-1.0 (1.9)	(-4.2, 2.2)	0.60	61	54	3.4 (2.3)	(-0.4, 7.2)	0.14
PY 3	47	41	-0.5 (1.9)	(-3.7, 2.6)	0.77	45	39	-2.6 (2.5)	(-6.8, 1.5)	0.29	49	43	2.0 (3.0)	(-2.8, 6.9)	0.49
PY 4	43	38	-0.2 (2.0)	(-3.6, 3.1)	0.92	41	35	-2.5 (2.6)	(-6.8, 1.8)	0.35	46	41	1.2 (3.3)	(-4.2, 6.5)	0.72
PY 5	38	33	-0.3 (2.2)	(-3.8, 3.3)	0.90	35	30	-2.7 (3.0)	(-7.7, 2.3)	0.37	41	36	1.5 (3.3)	(-4.0, 6.9)	0.66
PY 1 through 5	50	44	0.3 (1.4)	(-2.0, 2.7)	0.83	48	41	-1.4 (1.8)	(-4.3, 1.6)	0.44	52	47	1.9 (2.3)	(-1.8, 5.6)	0.41
Unweighted sample sizes															
Number of clinicians	7,590	28,324				3,825	16,302				3,808	12,234			
Number of clinician-years	26,257	94,698				13,197	54,976				13,060	39,722			

Source: Analysis of Medicare claims data from January 2014 through December 2021.

Notes: Although this table indicates which estimates are statistically significant, when we interpret evidence, we combine evidence from the magnitude of the effect, the *p*-values, findings on related outcomes, subgroups, sensitivity tests, and other data sources about model implementation.

[a] We report the actual, unadjusted averages in the baseline period, which are similar for the CPC+ and comparison groups due to matching. In the intervention periods, the comparison group mean is computed by subtracting the regression-adjusted difference between the CPC+ and comparison means in each time period from the CPC+ mean in that same time period.

[b] Each impact estimate is regression adjusted using a difference-in-differences analysis that reflects the difference between the average outcome for CPC+ clinicians in the five years of CPC+ and the average outcome in the baseline year, relative to the same difference over time for comparison clinicians, while controlling for prescriber's age, gender, and an indicator for whether the prescriber was a physician (versus a nurse practitioner or a physician's assistant), practice fixed effects, changes in state-level PDMP characteristics and opioid funding, county-level opioid marketing intensity, and COVID-19 controls for 2020 and 2021.

[c] After accounting for weights that adjust for matching, the effective sample sizes fall but are still substantial.

*/**/*** Significantly different from zero at the 0.10/0.05/0.01 level, two-tailed test.

C = comparison; CPC+ = Comprehensive Primary Care Plus; FFS = fee-for-service; MMEs = morphine milligram equivalents; n.a. = not applicable because the difference-in-differences impact estimate cannot be calculated at baseline; PDMP = prescription drug monitoring program; PY = Program Year; SE = standard error; SSP = Medicare Shared Savings Program.

Table 5.G.16. Difference-in-differences estimates of the effect of CPC+ on opioid prescribing over the five program years among beneficiaries who use opioids long term by SSP status, Track 2

	Overall					SSP					Non-SSP				
	CPC+ mean[a]	C mean[a]	Impact estimate[b] (SE)	90 percent confidence interval	*p*-value	CPC+ mean[a]	C mean[a]	Impact estimate[b] (SE)	90 percent confidence interval	*p*-value	CPC+ mean[a]	C mean[a]	Impact estimate[b] (SE)	90 percent confidence interval	*p*-value
Total MMEs per prescription															
Baseline	1,280	1,247	n.a.	n.a.	n.a.	1,283	1,184	n.a.	n.a.	n.a.	1,278	1,294	n.a.	n.a.	n.a.
PY 1	1,187	1,201	-47.2** (20.6)	(-81.0, -13.3)	0.02	1,191	1,179	-87.0*** (31.4)	(-138.6, -35.4)	0.01	1,184	1,216	-16.2 (27.2)	(-61.1, 28.6)	0.55
PY 2	1,095	1,109	-46.4** (22.5)	(-83.4, -9.3)	0.04	1,070	1,080	-108.7*** (34.7)	(-165.7, -51.7)	0.00	1,115	1,139	-8.2 (30.2)	(-57.9, 41.6)	0.79
PY 3	999	966	0.2 (26.7)	(-43.7, 44.1)	0.99	982	954	-71.1* (41.1)	(-138.8, -3.4)	0.08	1,011	980	47.2 (34.8)	(-10.1, 104.5)	0.18
PY 4	952	974	-54.5** (27.5)	(-99.6, -9.3)	0.05	949	958	-108.5** (42.6)	(-178.6, -38.5)	0.01	955	998	-27.0 (36.7)	(-87.4, 33.3)	0.46
PY 5	931	953	-55.1* (29.7)	(-103.9, -6.3)	0.06	910	959	-147.4*** (48.1)	(-226.4, -68.3)	0.00	946	955	7.0 (37.4)	(-54.4, 68.5)	0.85
PY 1 through 5	1,031	1,038	-39.9** (19.7)	(-72.3, -7.5)	0.04	1,021	1,021	-98.6*** (30.1)	(-148.1, -49.1)	0.00	1,039	1,055	-0.1 (25.5)	(-42.1, 41.9)	1.00
Number of days' supply per prescription															
Baseline	30.1	28.8	n.a.	n.a.	n.a.	30.5	28.7	n.a.	n.a.	n.a.	29.8	28.8	n.a.	n.a.	n.a.
PY 1	29.9	28.8	-0.2 (0.3)	(-0.6, 0.2)	0.38	30.1	28.5	-0.2 (0.4)	(-0.9, 0.5)	0.65	29.8	29.1	-0.2 (0.3)	(-0.8, 0.3)	0.45
PY 2	29.3	28.5	-0.5* (0.3)	(-0.9, 0.0)	0.07	29.4	28.6	-1.0** (0.4)	(-1.7, -0.3)	0.01	29.3	28.4	-0.1 (0.4)	(-0.7, 0.5)	0.84
PY 3	27.6	26.7	-0.4* (0.3)	(-0.9, 0.0)	0.10	27.5	27.0	-1.3*** (0.4)	(-2.0, -0.6)	0.00	27.6	26.5	0.2 (0.3)	(-0.4, 0.7)	0.61
PY 4	27.8	27.0	-0.6** (0.3)	(-1.0, -0.1)	0.04	27.9	27.2	-1.1** (0.5)	(-1.8, -0.3)	0.02	27.7	26.8	0.0 (0.3)	(-0.6, 0.5)	0.94
PY 5	27.5	27.1	-0.9*** (0.3)	(-1.4, -0.4)	0.00	27.4	27.4	-1.9*** (0.5)	(-2.7, -1.0)	0.00	27.7	26.9	-0.2 (0.4)	(-0.8, 0.5)	0.64
PY 1 through 5	28.4	27.6	-0.5** (0.2)	(-0.8, -0.1)	0.03	28.4	27.5	-0.9** (0.4)	(-1.5, -0.3)	0.01	28.4	27.5	-0.1 (0.3)	(-0.5, 0.4)	0.76
Number of prescriptions															
Baseline	60	55	n.a.	n.a.	n.a.	52	51	n.a.	n.a.	n.a.	65	57	n.a.	n.a.	n.a.
PY 1	53	50	-1.6* (0.9)	(-3.0, -0.2)	0.07	45	47	-3.0** (1.2)	(-4.9, -1.0)	0.01	59	52	-0.6 (1.2)	(-2.6, 1.4)	0.60
PY 2	49	48	-4.0*** (1.3)	(-6.2, -1.8)	0.00	42	45	-3.7** (1.8)	(-6.8, -0.7)	0.04	55	51	-3.9** (1.9)	(-7.0, -0.8)	0.04
PY 3	39	40	-5.2*** (1.5)	(-7.7, -2.7)	0.00	34	38	-4.3** (2.2)	(-7.9, -0.8)	0.05	43	41	-5.9*** (2.0)	(-9.3, -2.5)	0.00

Table 5.G.16. ***(continued)***

	Overall					SSP					Non-SSP				
	CPC+ mean[a]	C mean[a]	Impact estimate[b] (SE)	90 percent confidence interval	*p*-value	CPC+ mean[a]	C mean[a]	Impact estimate[b] (SE)	90 percent confidence interval	*p*-value	CPC+ mean[a]	C mean[a]	Impact estimate[b] (SE)	90 percent confidence interval	*p*-value
PY 4	37	37	-4.9*** (1.7)	(-7.7, -2.1)	0.00	33	34	-1.8 (2.4)	(-5.9, 2.2)	0.45	40	39	-7.1*** (2.4)	(-11.1, -3.2)	0.00
PY 5	32	33	-5.1*** (1.8)	(-8.1, -2.1)	0.00	29	31	-2.7 (2.5)	(-6.9, 1.5)	0.29	34	33	-6.8*** (2.4)	(-10.8, -2.7)	0.01
PY 1 through 5	42	41	-3.8*** (1.2)	(-5.7, -1.9)	0.00	37	39	-3.2** (1.6)	(-5.8, -0.6)	0.04	46	42	-4.2*** (1.6)	(-6.8, -1.6)	0.01
Unweighted sample sizes															
Number of clinicians	10,608	24,934				4,548	12,126				6,141	12,975			
Number of clinician-years	36,917	82,335				15,638	40,416				21,279	41,919			

Source: Analysis of Medicare claims data from January 2014 through December 2021.

Notes: Although this table indicates which estimates are statistically significant, when we interpret evidence, we combine evidence from the magnitude of the effect, the *p*-values, findings on related outcomes, subgroups, sensitivity tests, and other data sources about model implementation.

[a] We report the actual, unadjusted averages in the baseline period, which are similar for the CPC+ and comparison groups due to matching. In the intervention periods, the comparison group mean is computed by subtracting the regression-adjusted difference between the CPC+ and comparison means in each time period from the CPC+ mean in that same time period.

[b] Each impact estimate is regression adjusted using a difference-in-differences analysis that reflects the difference between the average outcome for CPC+ clinicians in the five years of CPC+ and the average outcome in the baseline year, relative to the same difference over time for comparison clinicians, while controlling for prescriber's age, gender, and an indicator for whether the prescriber was a physician (versus a nurse practitioner or a physician's assistant), practice fixed effects, changes in state-level PDMP characteristics and opioid funding, county-level opioid marketing intensity, and COVID-19 controls for 2020 and 2021.

[c] After accounting for weights that adjust for matching, the effective sample sizes fall but are still substantial.

*/**/*** Significantly different from zero at the 0.10/0.05/0.01 level, two-tailed test.

C = comparison; CPC+ = Comprehensive Primary Care Plus; MMEs = morphine milligram equivalents; n.a. = not applicable because the difference-in-differences impact estimate cannot be calculated at baseline; PDMP = prescription drug monitoring program; PY = Program Year; SE = standard error; SSP = Medicare Shared Savings Program

Table 5.G.17. Difference-in-differences estimates of the effect of CPC+ on opioid prescribing over the five program years, by track

	Track 1		Track 2	
	Impact estimate[a]	Percentage impact	Impact estimate[a]	Percentage impact
Total MMEs per prescription				
PY 1	-30.8	-2.5%	-47.2**	-3.8%
PY 2	-21.3	-1.9%	-46.4**	-4.1%
PY 3	-38.4	-3.8%	0.2	0.0%
PY 4	-51.1*	-5.1%	-54.5**	-5.4%
PY 5	-58.0**	-6.0%	-55.1*	-5.6%
PY 1 through 5	-37.4*	-3.5%	-39.9**	-3.7%
Number of days' supply per prescription				
PY 1	-0.3	-0.8%	-0.2	-0.7%
PY 2	-0.5*	-1.6%	-0.5*	-1.6%
PY 3	-0.5*	-1.8%	-0.4*	-1.6%
PY 4	-0.7**	-2.3%	-0.6**	-2.1%
PY 5	-0.7**	-2.5%	-0.9***	-3.2%
PY 1 through 5	-0.5**	-1.6%	-0.5**	-1.6%
Number of prescriptions				
PY 1	0.7	1.1%	-1.6*	-2.9%
PY 2	1.4	2.4%	-4.0***	-7.5%
PY 3	-0.5	-1.2%	-5.2***	-11.7%
PY 4	-0.2	-0.5%	-4.9***	-11.8%
PY 5	-0.3	-0.7%	-5.1***	-13.7%
PY 1 through 5	0.3	0.6%	-3.8***	-8.3%

Source: Analysis of Medicare claims data from January 2014 through December 2021.

Notes: Although this table indicates which estimates are statistically significant, when we interpret evidence, we combine evidence from the magnitude of the effect, the *p*-values, findings on related outcomes, subgroups, sensitivity tests, and other data sources about model implementation.

[a] Each impact estimate is regression adjusted using a difference-in-differences analysis that reflects the difference between the average outcome for CPC+ clinicians in the five years of CPC+ and the average outcome in the baseline year, relative to the same difference over time for comparison clinicians, while controlling for prescriber's age, gender, and an indicator for whether the prescriber was a physician (versus a nurse practitioner or a physician's assistant), practice fixed effects, changes in state-level PDMP characteristics and opioid funding, county-level opioid marketing intensity, and COVID-19 controls for 2020 and 2021. Impact estimates and standard errors are also shown in Tables 5.G.13 and 5.G.14.

*/**/*** Significantly different from zero at the 0.10/0.05/0.01 level, two-tailed test.

CPC+ = Comprehensive Primary Care Plus; MMEs = morphine milligram equivalents; PY = Program Year.

Table 5.G.18. Difference-in-differences (falsification test) results for the effect of CPC+ on opioid prescribing in 2016, by track and by SSP status, Track 1

	Overall					SSP					Non-SSP				
	CPC+ mean	C mean[a]	Impact estimate[b] (SE)	90 percent confidence interval	*p*-value	CPC+ mean	C mean[a]	Impact estimate[b] (SE)	90 percent confidence interval	*p*-value	CPC+ mean	C mean[a]	Impact estimate[b] (SE)	90 percent confidence interval	*p*-value
Total MMEs per prescription															
2015	1,328	1,253	n.a.	n.a.	n.a.	1,373	1,267	n.a.	n.a.	n.a.	1,283	1,238	n.a.	n.a.	n.a.
2016	1,283	1,199	8.2 (20.7)	(-32.4, 48.9)	0.69	1,335	1,220	8.5 (34.0)	(-58.2, 75.2)	0.80	1,229	1,174	9.1 (23.4)	(-36.7, 54.9)	0.70
Number of days' supply per prescription															
2015	30.0	28.8	n.a.	n.a.	n.a.	30.3	28.9	n.a.	n.a.	n.a.	29.7	28.7	n.a.	n.a.	n.a.
2016	30.2	28.8	0.2 (0.2)	(-0.3, 0.6)	0.48	30.6	29.0	0.2 (0.4)	(-0.5, 0.9)	0.51	29.8	28.7	0.1 (0.3)	(-0.5, 0.7)	0.74
Number of prescriptions															
2015	69	64	n.a.	n.a.	n.a.	67	59	n.a.	n.a.	n.a.	71	70	n.a.	n.a.	n.a.
2016	69	63	1.1 (0.9)	(-0.7, 3.0)	0.22	66	59	-0.3 (1.3)	(-2.8, 2.2)	0.79	72	68	2.7* (1.4)	(0.0, 5.4)	0.05
Unweighted sample sizes[c]															
Number of clinicians	4,672	17,371				2,368	9,935				2,305	7,444			
Clinician-years	8,592	31,391				4,340	17,928				4,252	13,463			

Source: Analysis of Medicare claims data from January 2013 through December 2016.

Notes: Although this table indicates which estimates are statistically significant, when we interpret evidence, we combine evidence from the magnitude of the effect, the p-values, findings on related outcomes, subgroups, sensitivity tests, and other data sources about model implementation.

[a] We report the actual, unadjusted averages in 2015. In 2016, the comparison group mean is computed by subtracting the regression adjusted difference between the CPC+ and comparison means from the CPC+ mean.

[b] Each impact estimate is regression adjusted using a difference-in-differences analysis that reflects the difference between the average outcome for Medicare FFS beneficiaries assigned to CPC+ practices in 2016 and the average outcome in 2015, relative to the same difference over time for Medicare FFS beneficiaries assigned to comparison practices, while controlling for beneficiary characteristics and practice fixed effects. We did not include changes in PDMP characteristics because we did not collect those data for 2015.

[c] After accounting for weights that adjust for matching, the effective sample sizes fall but are still substantial.

* Significantly different from zero at the 0.10 level, two-tailed test.

C = comparison; CPC+ = Comprehensive Primary Care Plus; FFS = fee-for-service; MMEs = morphine milligram equivalents; n.a. = not applicable because the difference-in-differences impact estimate cannot be calculated at baseline; SE = standard error; SSP = Medicare Shared Savings Program.

Table 5.G.19. Difference-in-differences (falsification test) results for the effect of CPC+ on opioid prescribing in 2016, by track and by SSP status, Track 2

	Overall					SSP					Non-SSP				
	CPC+ mean[a]	C mean[a]	Impact estimate[b] (SE)	90 percent confidence interval	*p*-value	CPC+ mean[a]	C mean[a]	Impact estimate[b] (SE)	90 percent confidence interval	*p*-value	CPC+ mean[a]	C mean[a]	Impact estimate[b] (SE)	90 percent confidence interval	*p*-value
Total MMEs per prescription															
2015	1,325	1,296	n.a.	n.a.	n.a.	1,328	1,231	n.a.	n.a.	n.a.	1,322	1,345	n.a.	n.a.	n.a.
2016	1,280	1,252	-0.5 (19.8)	(-39.2, 38.2)	0.98	1,283	1,182	4.0 (28.9)	(-52.6, 60.5)	0.89	1,278	1,304	-4.0 (26.9)	(-56.8, 48.8)	0.88
Number of days' supply per prescription															
2015	29.7	28.5	n.a.	n.a.	n.a.	30.1	28.3	n.a.	n.a.	n.a.	29.5	28.7	n.a.	n.a.	n.a.
2016	30.1	28.7	0.2 (0.2)	(-0.2, 0.7)	0.31	30.5	28.5	0.2 (0.4)	(-0.5, 0.9)	0.60	29.8	28.8	0.3 (0.3)	(-0.3, 0.9)	0.37
Number of prescriptions															
2015	60	55	n.a.	n.a.	n.a.	53	52	n.a.	n.a.	n.a.	66	57	n.a.	n.a.	n.a.
2016	60	55	-0.8 (0.8)	(-2.4, 0.7)	0.28	52	51	0.3 (1.0)	(-1.7, 2.4)	0.74	65	58	-1.7 (1.1)	(-3.9, 0.5)	0.13
Unweighted sample sizes[c]															
Number of clinicians	6,472	15,130				2,802	7,453				3,679	7,686			
Clinician-years	11,839	27,130				5,089	13,314				6,750	13,816			

Source: Analysis of Medicare claims data from January 2013 through December 2016.

Notes: Although this table indicates which estimates are statistically significant, when we interpret evidence, we combine evidence from the magnitude of the effect, the p-values, findings on related outcomes, subgroups, sensitivity tests, and other data sources about model implementation.

[a] We report the actual, unadjusted averages in 2015. In 2016, the comparison group mean is computed by subtracting the regression adjusted difference between the CPC+ and comparison means from the CPC+ mean.

[b] Each impact estimate is regression adjusted using a difference-in-differences analysis that reflects the difference between the average outcome for Medicare FFS beneficiaries assigned to CPC+ practices in 2016 and the average outcome in 2015, relative to the same difference over time for Medicare FFS beneficiaries assigned to comparison practices, while controlling for beneficiary characteristics, and practice fixed effects. We did not include changes in PDMP characteristics because we did not collect those data for 2015.

[c] After accounting for weights that adjust for matching, the effective sample sizes fall but are still substantial.

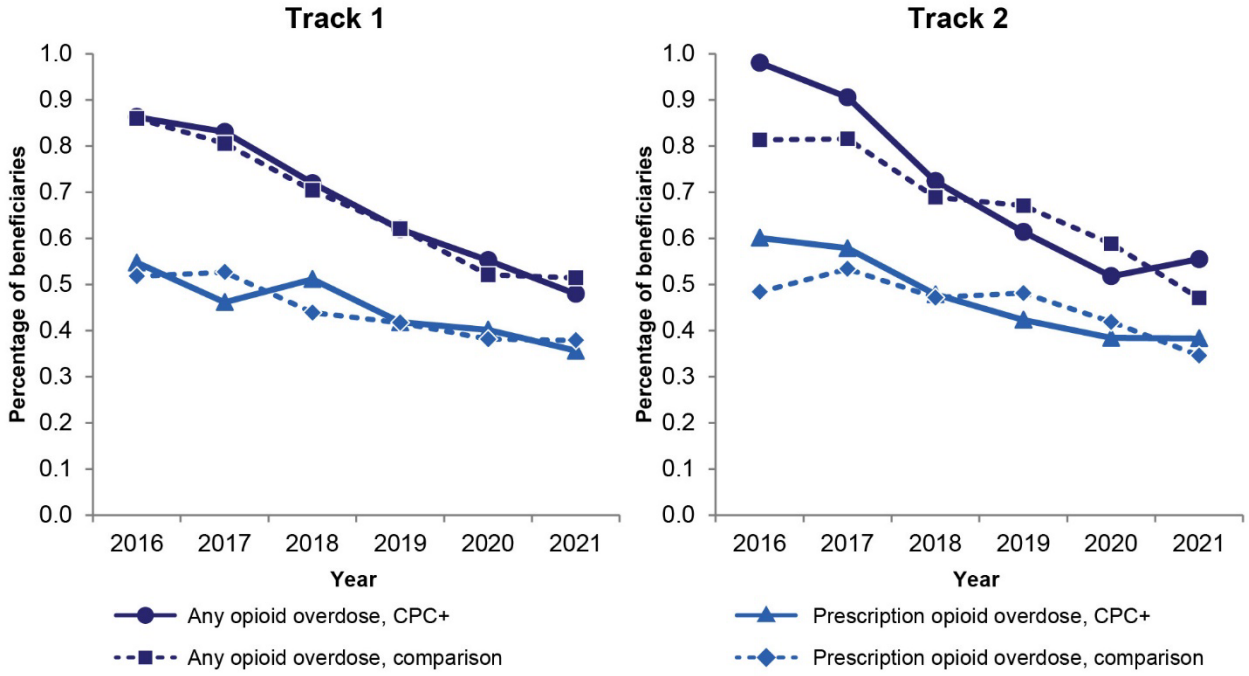
C = comparison; CPC+ = Comprehensive Primary Care Plus; MMEs = morphine milligram equivalents; FFS = fee-for-service; n.a. = not applicable because the difference-in-differences impact estimate cannot be calculated at baseline; SE = standard error; SSP = Medicare Shared Savings Program.

D. Trends in opioid overdoses

We found that less than one percent of beneficiaries who use opioids long term experienced an opioid overdose. Among Track 1 CPC+ and comparison beneficiaries who use opioids long term, the proportion who had an opioid overdose from either prescription or illicit opioids, declined from 0.9 percent in 2016 to 0.5 percent in 2021. Prescription overdoses comprised the largest proportion of all overdoses. The proportion of beneficiaries with a prescription overdose declined from 0.6 percent to 0.4 percent for CPC+ beneficiaries and from 0.5 percent to 0.4 percent for comparison beneficiaries over the same period (Figure 5.G.4, panel 1).

Change in overdoses over time was similar for Track 2 beneficiaries (Figure 5.G.4, panel 2).

Figure 5.G.4 Trends in any opioid overdose and prescription opioid overdoses among Medicare FFS beneficiaries who use opioids long term by CPC+ status, year, and track



CPC+ = Comprehensive Primary Care Plus; FFS = fee-for-service.

5.G.4. Discussion

Overall opioid prescribing and high-dose prescribing have been decreasing in the United States in the last several years (CDC 2019, 2020). This is consistent with our findings, which show declines in potential overuse in our sample between 2016 and 2021. The emergence of favorable effects on potential opioid overuse starting in the third year of CPC+ is consistent with the CPC+ theory of change that primary care transformation would require time to translate into improved outcomes. Both CPC+ and comparison prescribers reduced the number of prescriptions as well as the dose and duration of remaining prescriptions. CPC+ prescribers in both tracks reduced the dose and duration more than comparison prescribers, suggesting that reductions in CPC+ clinicians' prescribing drove improvements in potential opioid overuse. Further, we found that both overall and prescription opioid overdoses declined in CPC+ and comparison groups, but we could not analyze the impact of CPC+ on overdoses because of insufficient power.

The key limitation of this analysis is that we could not directly measure the mechanisms by which CPC+ reduced opioid prescribing and potential opioid overuse. We did not have longitudinal data on all relevant aspects of the CPC+ intervention that could have contributed to favorable estimated impacts. Also, several intervention components likely interacted to produce a favorable effect, and we did not believe we could aptly model these relationships.

That said, several aspects of the intervention may have contributed to the favorable findings. A previous study found that CPC+ practices increased their adoption of CMM and integrated behavioral health care between the first and the fourth year of the intervention. For example, the proportion of practices with access to an on-site pharmacist increased substantially: for Track 1, from 14 percent to 24 percent; for Track 2, from 21 to 52 percent (Swankoski et al. 2022). Further, Track 2 CPC+ prescribers also reduced the number of prescriptions relative to comparison prescribers, consistent with much greater implementation of CMM among Track 2 practices. Also, the proportion of CPC+ practices with access to a behaviorist more than doubled over the first four program years: for Track 1, from 18 percent to 45 percent; for Track 2, from 31 to 68 percent. Recent-small scale research studies have found implementation of similar interventions to be effective in reducing opioid prescribing in, for example, enhanced medication management (Parchman et al. 2019) and multimodal pain care intervention that included integration of behaviorists and pharmacists within primary care practices (Seal et al. 2020).

Several other aspects of the CPC+ intervention may have contributed to favorable findings; for example, improvements in care delivery in terms of access, continuity, comprehensiveness and coordination of care, and the use of data for planned care and population health. Within the area of comprehensiveness, for instance, practices were encouraged to screen patients for health-related social needs and to help address them through connections with community-based social service entities, a process identified as key in reducing reliance on opioids (Kerns 2022). Although we cannot conclude which intervention components drove findings, future primary care interventions that build on elements of the CPC+ model could more explicitly focus on opioid prescribing behaviors to investigate further.

Our analysis has additional limitations. First, despite having similar observed characteristics at baseline, CPC+ and comparison practices could differ on unobserved characteristics that may influence opioid use, particularly given that they are in different geographic areas. That said, our results were not sensitive to inclusion of different regional variables (defining opioid funding differently and adding pharmaceutical marketing controls). Second, with only two years of pre-intervention data on potential opioid overuse (two data points, given it is an annual outcome), we were unable to fully test for parallel trends in

outcomes between CPC+ and comparison practices. However, the falsification test results show similar changes for CPC+ and comparison groups in potential opioid overuse and prescribing outcomes from two years to one year before CPC+. Third, even though the potential opioid overuse outcome excludes most of the beneficiaries for whom such use is likely appropriate (those with cancer or sickle cell disease, and those who use hospice), it does not consider all appropriate use, such as use in non-hospice palliative care (CMS 2018). However, this would not affect the impact estimates because we measured potential opioid overuse consistently for CPC+ and comparison groups and we did not observe differential changes between these groups over time in a range of beneficiary characteristics over the duration of the program.

In conclusion, we found that CPC+, a large-scale intervention designed to transform the delivery of primary care, reduced potential opioid overuse among Medicare FFS beneficiaries. Although the magnitude of these impacts was modest, they translate into a substantial number of avoidable cases of opioid overuse. Further, even a small reduction in potential overuse is meaningful because this outcome captures use at a very high dosage for an extended period, which is associated with a high risk of serious adverse effects and overdose (Chou et al. 2020; Els et al. 2017; Von Korff et al. 2011).

5.H. Analysis of longer-term effects of CPC Classic

This Appendix examines the longer-term effects of primary care transformation—through the four-year Comprehensive Primary Care initiative (CPC Classic) and its five-year successor Comprehensive Primary Care Plus (CPC+)—on Medicare Part A and B expenditures (excluding enhanced payments made for CPC Classic, CPC+, or the Shared Savings Program [SSP]) and health care service use. In this Appendix, we first introduce the motivation for this analysis and the CPC Classic and CPC+ models (Section 1). We next explain the analytic methods (Section 2). Finally, we describe the results (Section 3) and discuss their implications (Sections 4 and 5).

5.H.1. Introduction

A. Background

Payers around the country are testing the patient-centered medical home (PCMH) and similar models and are increasingly paying for health care through alternative payment models that reward quality and value. Researchers and practitioners have warned that it takes time to transform care and shift patient outcomes (Nutting et al. 2009; Crabtree et al. 2011; McNellis et al. 2013; Peikes et al. 2020), but there have been no long-term models to assess whether the generally minimal changes that have been documented in outcomes such as emergency department (ED) visits and hospitalizations improve with longer interventions or follow-up periods. Against this backdrop, it is important to understand how longer tests of these models are associated with health care spending and utilization.

The Centers for Medicare & Medicaid Services (CMS) launched the four-year multipayer CPC Classic in October 2012. The goals of CPC Classic were to improve primary care delivery, health care quality, and patient experience, and to lower costs. CPC Classic also aimed to enhance clinicians' and staff members' experience. Across the country, 502 practices participated in CPC Classic; after CPC Classic ended in 2016, 85 percent of them immediately joined its five-year successor, CPC+, in 2017.

This analysis takes advantage of this unusually long combined model to examine the longer-term effects of primary care transformation on expenditures and service use for Medicare fee-for-service (FFS) beneficiaries. We examine effects over nine years—the four years of CPC Classic and five years after, in which many practices participated in the five-year successor, CPC+. We hypothesized that favorable effects of primary care transformation would emerge or remain over time.

B. Intervention

CPC Classic tested whether it was possible to reduce spending and improve quality by requiring primary care practices to improve care delivery in five areas: (1) access to and continuity of care, (2) planned care for preventive and chronic needs, (3) risk-stratified care management, (4) engagement of patients and their caregivers, and (5) coordination of care with patients' other care providers. The model provided substantially enhanced payments, including a $20 per beneficiary per month (PBPM) care management fee (CMF) from CMS in the first two years, and a $15 PBPM CMF in the last two years, as well as data feedback and learning support. A total of 502 primary care practices participated in CPC Classic.

Building on the lessons of CPC Classic and other advanced primary care models, in January 2017, CMS launched the five-year CPC+ model, which is the largest and most ambitious primary care payment and delivery reform ever tested in the United States (Anglin et al. 2020). Table 5.H.1 shows the main features of the two models were similar, with the notable differences being that CPC+:

- Was larger in size,
- Increased the emphasis on aspects of comprehensiveness, including behavioral health integration and assessing and addressing patients' social support needs,
- Allowed simultaneous participation in the Medicare Shared Savings Program (SSP),
- Included a more advanced care transformation track with the following features:
  - Health information technology vendor support,
  - Substantially higher enhanced payments and progressively larger alternative-to-FFS payments, and
  - Requirements for some more advanced care delivery approaches.

CMS offered all CPC Classic practices participation in CPC+ if they met basic eligibility criteria. After CPC Classic ended, many of the CPC Classic practices (85 percent) joined CPC+ in 2017, predominantly in Track 2. Specifically, 71 CPC Classic practices joined Track 1 of CPC+ and constituted 5 percent of all Track 1 practices that began CPC+ in 2017; 352 CPC Classic practices joined Track 2 of CPC+ and constituted 23 percent of all Track 2 2017 Starters in CPC+.

Table 5.H.1. Comparison of CPC Classic and CPC+

	CPC Classic	CPC+
Model		
Model duration	Four years (October 2012–December 2016)	Five years (January 2017–December 2021)
Care delivery requirements	(1) Access to and continuity of care, (2) planned care for preventive and chronic needs, (3) risk-stratified care management, (4) engagement of patients and their caregivers, and (5) coordination of care with patients' other care providers	(1) Access and continuity, (2) care management, (3) comprehensiveness and coordination, (4) patient and caregiver engagement, and (5) planned care and population health. CPC+ increased the emphasis on aspects of comprehensiveness, including behavioral health integration and assessing and addressing patients' social support needs. CPC+ includes two tracks with different levels of care delivery requirements and payment approaches to meet the diverse needs of participating practices. Track 2 practices are required to provide more enhanced care delivery approaches to better support patients with complex needs than Track 1 practices, and they receive higher payments.

Table 5.H.1. ***(continued)***

	CPC Classic	CPC+
Reach		
Partners	CMS 39 other private and public payers	CMS 79 other private and public payers 68 health IT vendors
Number of regions	7	18
Number of intervention practices	502	3,070 (1,504 in Track 1 and 1,566 in Track 2)
Number of beneficiaries served	Over 2.5 million	Over 17 million
Supports		
Average of risk-adjusted care management fees PBPM[a]	From CMS: $20 in first two years, $15 in last two years; lower from other payers	From CMS: $15 for Track 1, $28 PBPM for Track 2; lower from other payers
Median enhanced funding per practice (also calculated per primary care practitioner) in the latest model year (4 for CPC, and 2 for CPC+)[b, c]	$179,519 (or $50,189 per practitioner), or 10 percent of practice revenue	Track 1: $122,065 (or $42,964 per practitioner), or 10 percent of practice revenue Track 2: $263,606 (or $66,424 per practitioner), or 15 percent of practice revenue
Payments other than CMFs[b]	Share in any savings after covering CMFs starting in Year 2, offered by Medicare FFS and two-thirds of other payers.	Payments for performance on cost, utilization, and/or quality measures, offered by Medicare FFS and 94 percent of other payers. Unlike CPC Classic, CPC+ practices also have the option to participate in Medicare SSP. If they do, they can earn shared savings from that program but are not eligible for performance-based payments from CPC+ because of CMS's rules that prohibit "double dipping". Alternative to FFS payments starting in CPC+ Year 1 by CMS and 22 percent of payer partners in Year 2 for Track 2. A portion of FFS payments was converted to lump sum payments regardless of visits.
Non-financial supports	Data feedback, learning support	Data feedback, learning support, and health IT vendor support

[a] CMS risk adjusts CMFs based on beneficiaries' hierarchical condition category score, which is a claims-based measure of risk for subsequent expenditures.

[b] Numbers reported in the CPC+ column apply to all practices that joined CPC+ in 2017 and are not limited to the CPC Classic alumni.

[c] The enhanced funding included CMFs and performance-based payments. In Year 2 of CPC+, CMFs represented 90 percent of total enhanced funding.

CMFs = care management fees; CMS = Centers for Medicare & Medicaid Services; CPC Classic = Comprehensive Primary Care initiative; CPC+ = Comprehensive Primary Care Plus; FFS = fee-for-service; IT = information technology; PBPM = per beneficiary per month, SSP = Medicare Shared Savings Program.

5.H.2. Methods

A. Evaluation design

To measure the effects of primary care transformation with service use and spending, we compared changes in outcomes from the year before CPC Classic began (baseline period) to the nine-year period after it began (intervention period), between Medicare FFS beneficiaries served by intervention practices (defined as those that began CPC Classic and were still participating at the end of the first quarter) and those served by matched comparison practices. We used propensity score matching to ensure pre-intervention similarity between intervention and comparison practices across beneficiary, practice, and market characteristics. Matching variables included beneficiaries' characteristics (such as age, sex, HCC scores, and prior expenditures and service use); practice-level characteristics (such as meaningful use of electronic health records, number of clinicians, and percentage of clinicians with a primary care specialty); and characteristics of the practice's market (such as mean county income). We selected as many as five comparison practices for each CPC Classic practice.

Starting in the first quarter of CPC Classic through the fourth intervention year (October 2012 to December 2016), Medicare FFS beneficiaries were attributed quarterly to CPC and comparison practices that delivered the largest share of their primary care visits during a two-year lookback period. We then used an intent-to-treat (ITT) design to assign beneficiaries to practices in the intervention period; that is, once we had attributed beneficiaries to a practice (intervention or comparison) at any time during the intervention period, they remained in the analysis sample as long as they met the eligibility criteria (alive and enrolled in Medicare Part A and Part B with Medicare as the primary payer and not in a Medicare Advantage Plan). For the five years after CPC Classic ended (January 2017 to December 2021), we followed the beneficiaries already assigned in the fourth-year analysis sample into their fifth, sixth, seventh, eighth, and ninth years, with the same intervention or comparison status as in CPC Classic.

For the baseline period, we defined the study sample as beneficiaries who were attributed to the intervention or comparison practices during the intervention period and were alive at the start of the period. As a result, the baseline sample did not include people who had died during the baseline year. This meant that Medicare expenditures and service use during the baseline period were lower (for both the intervention and comparison groups) than in later periods because the baseline period did not include beneficiaries who needed expensive end-of-life care.

For details on matching methods, attribution, and ITT design, please refer to the supplemental appendix in Dale et al. (2016).

B. Measures of spending and utilization

We constructed three main outcomes from Medicare claims and enrollment data: (1) Medicare Part A and Part B expenditures excluding enhanced payments made for CPC Classic, CPC+, or SSP; (2) acute hospitalizations; and (3) outpatient ED visits. We also examined impacts on expenditures by service category: (1) inpatient (overall, acute, and non-acute), (2) outpatient, (3) physician, (4) home health, (5) hospice, (6) skilled nursing facility, and (7) durable medical equipment.

C. Statistical analysis

We implemented a difference-in-differences model that compares the mean change in outcomes from the year before the start of CPC Classic to the nine years after between two groups: (1) beneficiaries served by the CPC Classic practices and (2) beneficiaries served by comparison practices. We used (1) linear regressions for Medicare Part A and Part B expenditures and (2) zero-inflated negative binomial regressions for acute hospitalizations and outpatient ED visits to account for a large percentage of zeroes. The regressions controlled for beneficiary, practice, and market characteristics observed at baseline to net out observable pre-existing baseline differences between CPC Classic and comparison beneficiaries that remained after propensity score matching.

Since the analysis includes calendar years 2020 and 2021 (or Years 8 and 9), it was important to account for any differences in how the coronavirus disease 2019 (COVID-19) pandemic unfolded for the intervention and comparison practices. Therefore, we included the same COVID-19-related regional control variables in these years that were used in the main impact analysis from the CPC+ final report (see Appendix 5.E for details).

Estimated standard errors accounted for beneficiary outcomes clustered at the practice level and for weighting. The overall weights were equal to the product of two separate weights that accounted for (1) the share of the year for which the beneficiary's data were observed and (2) a matching weight (derived from the propensity score matching procedure) ensuring that CPC Classic and comparison practices were balanced.

We performed all statistical analyses with Stata software (Version 15.1). We provide p-values for all estimates and consider p-value < 0.10 to be statistically significant.

5.H.3. Results

A. Practices included in the study sample

The analysis included 497 practices participating at the end of CPC Classic's first quarter and 908 similar comparison practices. None of the comparison practices joined CPC Classic (by design); 21 percent joined CPC+ in 2017. The intervention and comparison groups had similar practice characteristics at baseline (Table 5.H.2) and similar trajectories of key outcomes (Medicare expenditures, hospitalizations, and outpatient ED visits) in the two years before CPC Classic began (Fu et al. 2022).[116]

[116] We analyzed acute hospitalizations as defined under the CPC+ evaluation, which is different from total hospitalizations included in Fu et al. (2022), for two reasons: (1) acute hospitalizations are more likely to be directly affected by the care delivery features of the intervention, and (2) the impact estimates on acute hospitalizations are directly comparable to those on acute inpatient expenditures in this analysis. Acute hospitalization claims constitute over 90 percent of all inpatient claims.

Table 5.H.2. Baseline practice characteristics of CPC Classic and comparison practices[a]

Characteristic	Intervention practices	Comparison practices	Difference between intervention and comparison practices	*p*-Value
Percentage of practices with one or more clinicians who was a Medicare meaningful EHR user as of June 2012[b]	79	79	0	>0.99
Percentage of practices with state or NCQA medical-home recognition by autumn 2012[c]	39	37	2.9	0.20
Mean number of clinicians[d]	4.2	4.6	-0.4	0.64
Percentage of practices' clinicians with primary care specialty	94	94	0	0.92
Percentage of practices owned by larger organization[d]	55	54	1	0.85
Percentage of practices located in medically underserved area[e]	11	14	-3	0.17
Percentage of practice's county that is urban[f]	78	75	3	0.08
Mean number of attributed Medicare beneficiaries[g]	635	698	-63	0.14

[a] Because the CPC Classic intervention was provided at the practice level, and to aid computation, we matched using practice-level data rather than beneficiary-level data. The means (rounded to whole numbers) in this table represent practice-level means, weighted to account for matching.

[b] A meaningful EHR user is a clinician who qualified for CMS incentive programs by having used certified EHR technology to improve the quality of health care and to meet other objectives specified by CMS.

[c] Numbers are based on September 2012 data from NCQA.

[d] Data are from a 2012 office-based physician file from SK&A, a health care marketing vendor.

[e] Numbers are based on 2009 data from the HRSA.

[f] Data are from the 2009 Area Health Resource Files provided by the HRSA.

[g] Numbers are based on 2010-2012 Medicare claims and enrollment data from the CMS Virtual Research Data Center.

CMS = Centers for Medicare & Medicaid Services; CPC = comprehensive primary care; EHR = electronic health record; HRSA = Health Resources and Services Administration; NCQA = National Committee for Quality Assurance.

B. Beneficiaries included in the study sample

The analysis included over 500,000 beneficiaries in the intervention group and over 1.1 million beneficiaries in the comparison group. Table 5.H.3 shows that the baseline beneficiary characteristics and outcomes for the intervention and comparison groups were similar.

Table 5.H.3. Baseline outcomes and characteristics of CPC Classic and comparison beneficiaries in the research sample[a]

Panel A. Baseline characteristics of beneficiaries included in the research sample[b]

Measure	Intervention mean[c] (N = 565,674)	Comparison mean[c] (N = 1,165,284)	Intervention-comparison difference	Standardized difference
Age				
Younger than 50	6.1	6.7	-0.6	-0.03
50–64	16.7	16.8	-0.2	0.00
65–74	41.2	41.0	0.2	0.01
75–84	24.8	24.8	0.0	0.00
85 or older	11.2	10.7	0.6	0.02
Race				
White	90.6	91.0	-0.4	-0.02
Black	4.4	4.5	-0.2	-0.01
Native American	1.8	1.1	0.7	0.06
Other	3.3	3.4	-0.1	-0.01
Male	41.7	42.1	-0.4	-0.01
Original reason for Medicare eligibility				
Age	78.5	77.3	1.2	0.03
Disabled	21.3	22.6	-1.2	-0.03
ESRD	0.1	0.1	0.0	0.00
Dually eligible for Medicare and Medicaid	11.4	13.1	-1.7	-0.06
HCC score (continuous measure)[d]	1.0	1.0	0.0	-0.01
HCC score originally missing and imputed	9.7	9.6	0.2	0.01

Panel B. Baseline outcomes of beneficiaries in the research sample who had baseline data

Measure	Intervention mean[c] (N = 442,709)	Comparison mean[c] (N = 954,199)	Intervention-comparison difference	Standardized difference
Main outcomes				
Medicare expenditures without fees (PBPM)	$574.1	$578.3	-$4.1	0.00
Acute hospitalizations (per 1,000 beneficiaries per year)	224.5	225.7	-1.2	0.00
Outpatient ED visits (per 1,000 beneficiaries per year)	417.4	440.5	-23.2	-0.02
Other outcomes: Expenditures by service category (PBPM)				
Inpatient	$204.8	$200.8	$3.9	0.01
Acute inpatient	$177.8	$177.7	$0.1	0.00
Non-acute inpatient	$27.0	$23.1	$3.8	0.01
Outpatient	$97.2	$103.1	-$5.8	-0.02
Physician	$199.6	$195.0	$4.6	0.01
Skilled nursing	$29.6	$31.8	-$2.3	-0.01
Home health	$26.3	$30.3	-$4.0	-0.04
Hospice	$2.0	$2.4	-$0.5	-0.01
Durable medical equipment	$22.5	$23.2	-$0.7	-0.01

Table 5.H.3. ***(continued)***

[a] Medicare claims and enrollment data for October 2011 through December 2021. The baseline outcomes are not available for beneficiaries who were added to the sample in later years but were not eligible at baseline. However, we were able to obtain the baseline characteristics for these beneficiaries using the following approach: (1) for race, gender, and original reason for Medicare eligibility at baseline, we used data from the time the beneficiary first became eligible; (2) we calculated age using the date of birth reported; (3) for dual eligibility, we conservatively assumed that these beneficiaries were not dual eligible at baseline; (4) for HCC scores, we imputed the baseline (2011) scores for these beneficiaries, specifically by using the average (non-missing) HCC score of 66-year-old beneficiaries for beneficiaries with missing HCC scores who were 65 years or older and the average (non-missing) HCC scores for beneficiaries below age 65 for beneficiaries with missing HCC scores who were under age 65.

[b] Data are percentages in Panel A, unless noted.

[c] Means (rounded to one decimal place) were weighted to account for (1) the share of the year for which the beneficiary's data were observed and (2) the matching (for beneficiaries in comparison practices only).

[d] HCC scores are a measure of risk for subsequent expenditures. CMS calculates them such that the average for the Medicare FFS population nationally is 1.0. A patient with a risk score of 1.30 is predicted to have expenditures that would be approximately 30 percent above the average, whereas a patient with a risk score of 0.70 is expected to have expenditures that would be approximately 30 percent below the average.

CMS = Centers for Medicare & Medicaid Services; ED = emergency department; ESRD = end-stage renal disease; FFS = fee-for-service; HCC = hierarchical condition category; PBPM = per beneficiary per month.

C. Difference-in-differences estimates for main outcomes

During the nine years since CPC Classic began, the cumulative estimates indicate that intervention and comparison practices had similar Medicare FFS expenditures over time. However, there was an overall slower growth in acute hospitalizations and outpatient ED visits among intervention practices, relative to comparison practices (Table 5.H.4).

When assessing the estimates during the four years of CPC Classic (2012–2016, or Years 1–4) and five years after (2017–2021, or Years 5–9, hereafter referred to as "post-Classic period") as well as the annual estimates (shown in Table 5.H.4, Figure 5.H.1, and Figure 5.H.2), we found the following:

1. Relative to comparison practices, beneficiaries in intervention practices experienced the following effects:
   - A relative reduction in acute hospitalizations (2.2 percent, $p = 0.02$) over the nine years after CPC Classic began, with no discernible differences between the intervention and comparison practices during the Classic period and a relative decrease of 3.0 percent during the post-Classic period ($p = 0.01$). The relative reduction in acute hospitalizations for CPC Classic beneficiaries was statistically significant in Year 2 (2 percent, $p = 0.09$), Year 5 (3.2 percent, $p = 0.01$), Year 6 (4.0 percent, $p < 0.01$), and Year 7 (4.1 percent, $p < 0.01$).
   - A relative reduction in outpatient ED visits (1.8 percent, $p = 0.09$) over the nine years after CPC Classic began, with a relative reduction during the CPC Classic period (1.6 percent, $p = 0.07$) and a somewhat larger but not statistically significant decrease during the post-Classic period (1.9 percent, $p = 0.15$). Relative to comparison beneficiaries, there was a statistically significant reduction in outpatient ED visits for CPC Classic beneficiaries in Year 3 (2.6 percent, $p = 0.01$), Year 4 (2.2 percent, $p = 0.04$), Year 5 (2.2 percent, $p = 0.09$), and Year 7 (3.9 percent, $p = 0.01$).
2. There was no discernible effect on CPC Classic beneficiaries' Medicare Part A and B expenditures excluding additional payments from CPC Classic, CPC+, or SSP in the nine years after CPC Classic began, during the CPC Classic period, or during the post-Classic period, relative to comparison beneficiaries. A statistically significant reduction in growth of expenditures (2.2 percent, $p = 0.01$) was observed in Year 1; however, it was too soon after the start of CPC Classic to be plausible as a causal impact. This reduction was not seen in any subsequent years.

Table 5.H.4. Regression-adjusted means and difference-in-differences estimates for service use and expenditures among attributed Medicare fee-for-service beneficiaries for CPC Classic and comparison practices, annual and nine-year cumulative estimates

	Intervention mean	Comparison mean	Difference-in-differences estimate (SE)	Difference-in-differences estimate in percentage[a]	90 percent confidence interval	*p*-Value
Service use (per 1,000 beneficiaries per year)						
Acute hospitalizations						
Baseline	224	226	NA	NA	NA	NA
Y1	317	322	-4.1 (3.4)	-1.3%	(-9.6, 1.5)	0.23
Y2	310	318	-6.3* (3.7)	-2.0%	(-12.4, -0.2)	0.09
Y3	315	318	-1.6 (3.7)	-0.5%	(-7.6, 4.4)	0.66
Y4	298	304	-4.5 (3.6)	-1.5%	(-10.5, 1.5)	0.22
Y5	292	303	-9.7*** (3.6)	-3.2%	(-15.7, -3.8)	0.01
Y6	293	307	-12.2*** (3.9)	-4.0%	(-18.7, -5.8)	0.00
Y7	298	312	-12.6*** (4.5)	-4.1%	(-19.9, -5.3)	0.00
Y8	261	269	-6.8 (4.5)	-2.5%	(-14.1, 0.6)	0.13
Y9	270	276	-4.4 (4.8)	-1.6%	(-12.2, 3.4)	0.35
Y1–Y4 (CPC Classic period)	310	315	-4.4 (3.1)	-1.4%	(-9.6, 0.8)	0.16
Y5–Y9 (Post-Classic period)	285	295	-8.9*** (3.2)	-3.0%	(-14.2, -3.6)	0.01
Y1–Y9	296	304	-6.8** (3.0)	-2.2%**	(-11.7, -1.9)	0.02
Outpatient ED visits, including observation stays						
Baseline	417	441	NA	NA	NA	NA
Y1	466	492	-2.0 (4.7)	-0.4%	(-9.7, 5.7)	0.68
Y2	489	515	-3.1 (5.1)	-0.6%	(-11.5, 5.3)	0.55
Y3	503	540	-13.4** (5.4)	-2.6%	(-22.3, -4.5)	0.01
Y4	502	537	-11.5** (5.7)	-2.2%	(-20.9, -2.1)	0.04
Y5	514	549	-11.8* (6.9)	-2.2%	(-23.1, -0.4)	0.09
Y6	515	547	-8.6 (7.0)	-1.6%	(-20.0, 2.9)	0.22
Y7	519	563	-21.0*** (7.8)	-3.9%	(-33.9, -8.1)	0.01

Table 5.H.4. ***(continued)***

	Intervention mean	Comparison mean	Difference-in-differences estimate (SE)	Difference-in-differences estimate in percentage[a]	90 percent confidence interval	*p*-Value
Y8	413	441	-4.7 (8.6)	-1.1%	(-18.9, 9.5)	0.59
Y9	455	481	-2.2 (8.9)	-0.5%	(-16.8, 12.5)	0.81
Y1–Y4 (CPC Classic period)	492	523	-8.2* (4.5)	-1.6%	(-15.6, -0.9)	0.07
Y5–Y9 (Post-Classic period)	488	520	-9.5 (6.5)	-1.9%	(-20.2, 1.2)	0.15
Y1–Y9	490	522	-8.9* (5.2)	-1.8%	(-17.5, -0.3)	0.09
Medicare expenditures (PBPM)						
Medicare Part A and B expenditures excluding enhanced payments made for CPC Classic, CPC+, or SSP						
Baseline	$574	$578	NA	NA	NA	NA
Y1	$774	$796	-$17.8*** ($6.6)	-2.2%	(-$28.6, -$6.9)	0.01
Y2	$802	$817	-$10.5 ($6.9)	-1.3%	(-$21.9, $0.9)	0.13
Y3	$837	$845	-$3.4 ($7.6)	-0.4%	(-$15.9, $9.0)	0.65
Y4	$857	$862	-$1.2 ($8.4)	-0.1%	(-$15.0, $12.5)	0.88
Y5	$905	$915	-$6.3 ($8.4)	-0.7%	(-$20.1, $7.6)	0.46
Y6	$946	$955	-$5.2 ($9.5)	-0.5%	(-$20.8, $10.5)	0.59
Y7	$1,021	$1,024	$1.2 ($10.2)	0.1%	(-$15.5, $18.0)	0.90
Y8	$986	$990	$0.2 ($10.5)	0.0%	(-$17.0, $17.4)	0.98
Y9	$1,124	$1,114	$14.3 ($12.7)	1.3%	(-$6.7, $35.2)	0.26
Y1–Y4 (CPC Classic period)	$821	$833	-$7.6 ($6.2)	-0.9%	(-$17.8, $2.6)	0.22
Y5–Y9 (Post-Classic period)	$987	$991	-$0.1 ($8.6)	0.0%	(-$14.2, $14.0)	0.99
Y1–Y9	$909	$917	-$3.6 ($7.0)	-0.4%	(-$15.2, $8.0)	0.61
Sample sizes						
Number of practices	497	908				
Number of beneficiaries	565,674	1,165,284				
Number of beneficiary years	3,955,515	8,118,902				

Source: Medicare claims data for October 2011 through December 2021.

Table 5.H.4. ***(continued)***

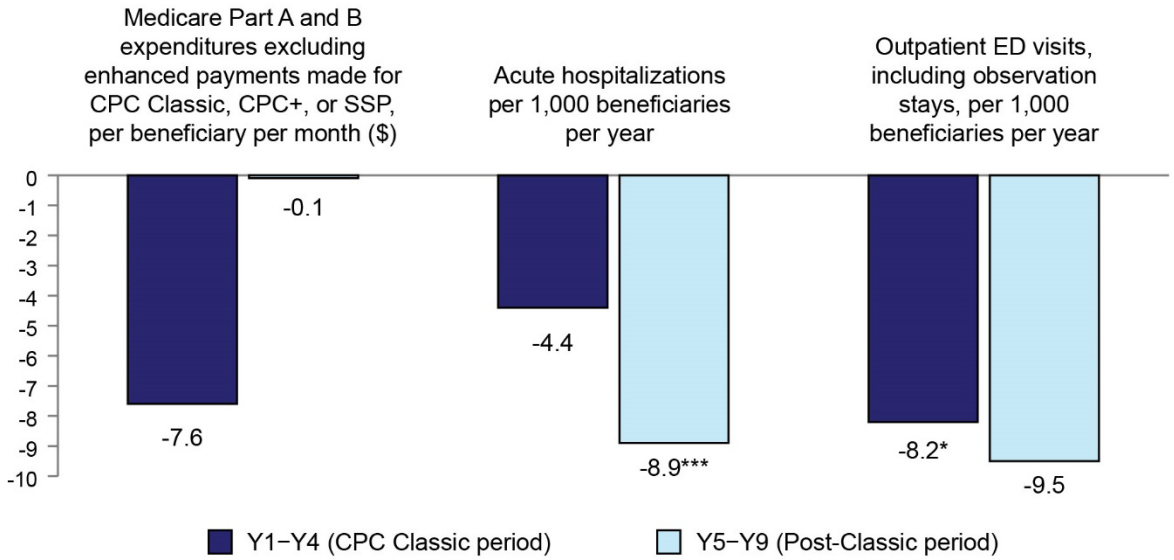
Notes: Estimates are regression-adjusted for baseline beneficiary characteristics, baseline practice characteristics, and COVID-19-related controls. We based each estimate on a difference-in-differences analysis, and each reflects the difference in the regression-adjusted average outcome for attributed Medicare FFS beneficiaries in intervention practices in Years 1 to 9 compared with baseline relative to the same difference over time for attributed Medicare FFS beneficiaries in comparison practices. Note that expenditures and utilization are generally lower in the baseline year (relative to intervention years) because the baseline sample was composed of beneficiaries who were attributed (and hence alive) during the intervention period and did not include beneficiaries who needed expensive end-of-life care who would have died during the baseline year.

*/**/*** Significantly different from zero at the 0.10/0.05/0.01 levels, two-tailed test.

[a] To calculate these percentages, we divided the difference-in-differences estimate by the mean for the outcome in the intervention group minus the difference-in-differences estimate.

COVID-19 = coronavirus disease 2019; CPC = Comprehensive Primary Care; CPC+ = Comprehensive Primary Care Plus; ED = emergency department; FFS = fee-for-service; NA = not applicable; PBPM = per beneficiary per month; SSP = Medicare Shared Savings Program; SE = standard error; Y = intervention year.

Figure 5.H.1. Estimated effects on expenditures and service use for attributed Medicare fee-for-service beneficiaries for CPC Classic and comparison practices, during CPC Classic period and post-Classic period



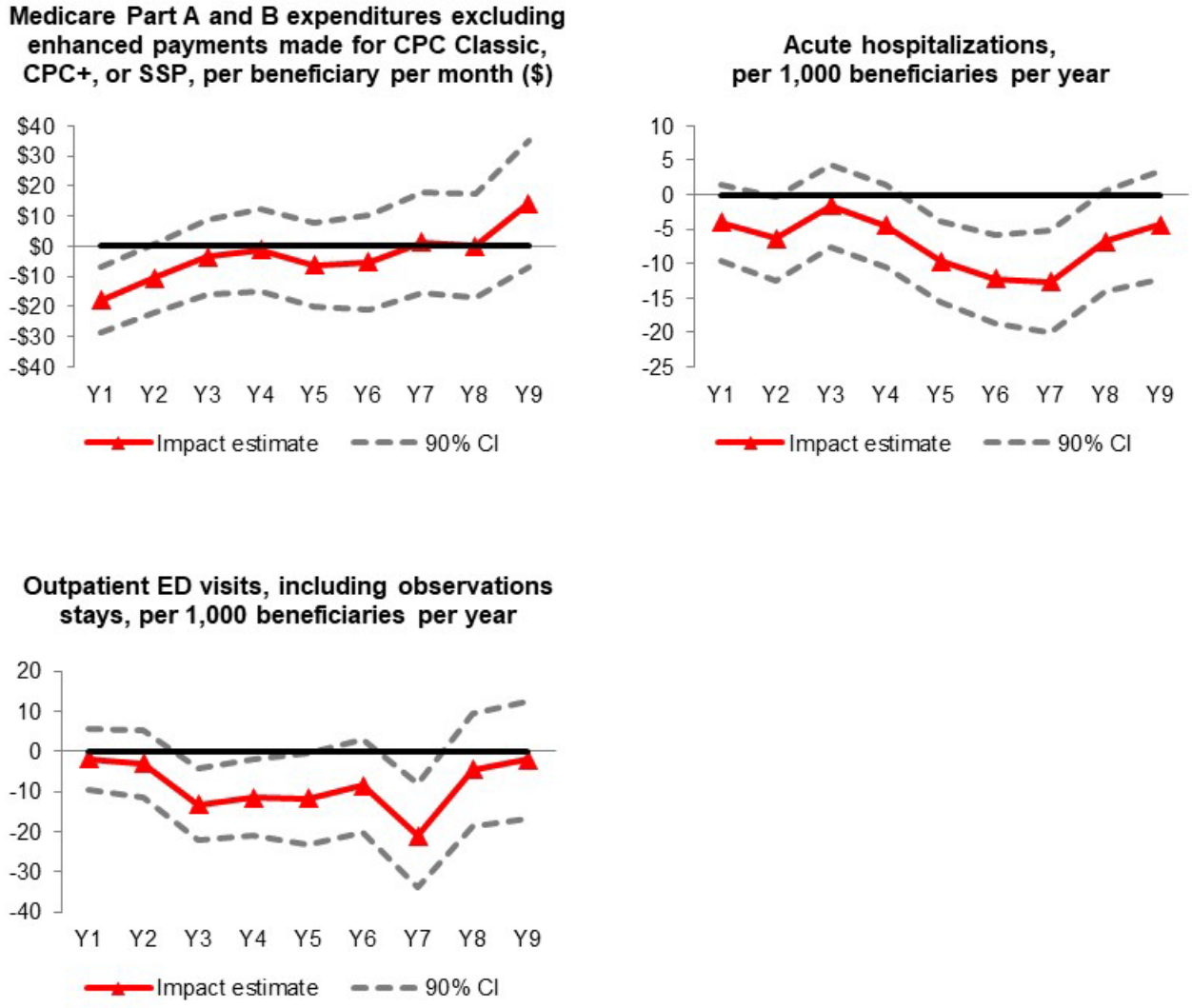
Source: Medicare claims data for October 2011 through December 2021.

Notes: The estimate of the effect is equal to the difference in mean outcomes between attributed Medicare FFS beneficiaries in the intervention and comparison group practices in the intervention period (that is, CPC Classic period or post-Classic period) minus the average difference between the two groups during the baseline period. Estimates are regression-adjusted for baseline beneficiary characteristics, baseline practice characteristics, and COVID-19-related controls.

*/**/*** Significantly different from zero at the 0.10/0.05/0.01 levels, two-tailed test.

COVID-19 = coronavirus disease 2019; CPC = Comprehensive Primary Care; CPC+ = Comprehensive Primary Care Plus; ED = emergency department; FFS = fee-for-service; SSP = Medicare Shared Savings Program, Y = intervention year.

Figure 5.H.2. Estimated effects on expenditures and service use for attributed Medicare fee-for-service beneficiaries for CPC Classic and comparison practices, by year



Source: Medicare claims data for October 2011 through December 2021.

Notes: The estimate of the effect, denoted by a separate triangle for each intervention year in the figure, is equal to the difference in mean outcomes between attributed Medicare FFS beneficiaries in the intervention and comparison group practices in any year since CPC Classic began minus the average difference between the two groups during the baseline period. Estimates are regression-adjusted for baseline beneficiary characteristics, baseline practice characteristics, and COVID-19-related controls. The dashed lines indicate the 90 percent confidence interval.

CI = confidence interval; COVID-19 = coronavirus disease 2019; CPC = Comprehensive Primary Care; CPC+ = Comprehensive Primary Care Plus; ED = emergency department; FFS = fee-for-service; SSP = Medicare Shared Savings Program; Y = intervention year.

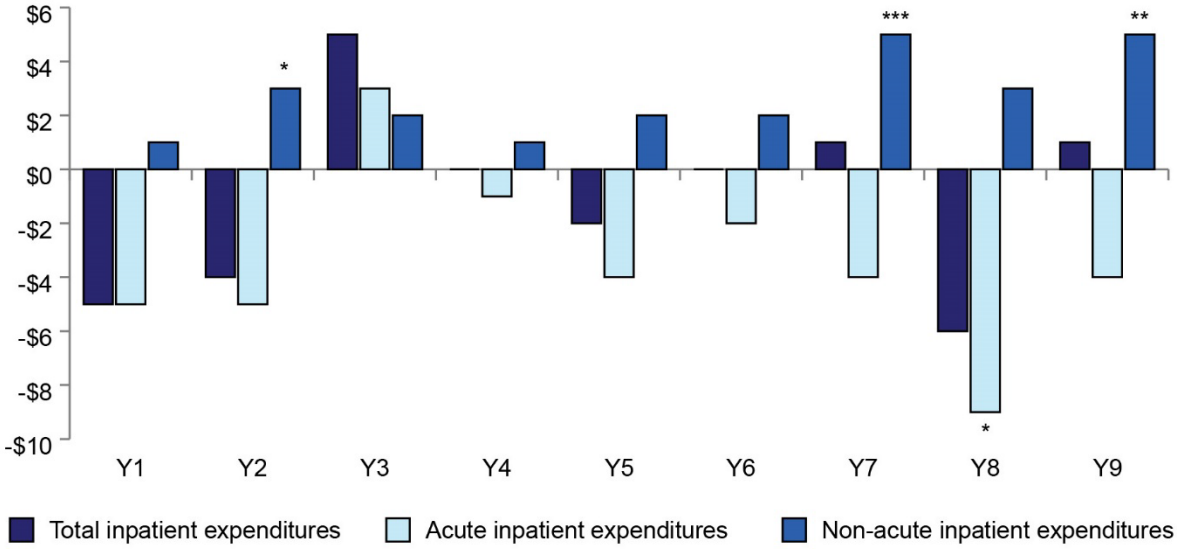
D. Difference-in-differences estimates for expenditures by service category

To try to understand why reductions in acute hospitalizations and ED visits did not translate into reduction in Medicare expenditures, we next examined the effects on specific expenditure categories. Note that expenditures in all categories increased over time for both CPC and comparison practices before the COVID-19 pandemic started in 2020. We use "reduction(s)" in expenditures as a shorthand below to describe "slower growth" or "relative reductions". We were particularly interested in inpatient and outpatient expenditures because they are most likely to be directly affected by the lower growth in acute hospitalizations. This analysis showed that there was no impact on inpatient expenditures, in spite of the impacts on acute hospitalizations. Also, while outpatient expenditures fell, the size of the reduction was only \$5.4 PBPM ($p < 0.01$), which was offset by small increases in other expenditure categories, such as hospice (Table 5.H.5).

Our specific findings include:

- There was no overall effect on inpatient expenditures across the nine years. When segregating inpatient expenditures into acute and non-acute inpatient expenditures, we found that in many years (except for Year 3), the relative decrease in acute inpatient expenditures (statistically significant in Year 8, with a reduction of 3.1 percent, or \$9.2 PBPM) were partially offset by a relative increase in non-acute inpatient expenditures (statistically significant in Years 2, 7, and 9, ranging from 7 percent to 12.2 percent, or \$2.7 to \$5.3 PBPM). As a result, there were no discernable differences in total inpatient expenditures between intervention and comparison practices (Figure 5.H.3).
- Over the nine years, there was a 3.2 percent (\$5.4 PBPM, $p < 0.01$) reduction in outpatient expenditures. The reductions in outpatient expenditures were statistically significant in both the CPC Classic and the post-Classic periods (\$3.1 PBPM, 2.3 percent, $p = 0.03$ and \$7.4 PBPM, 3.9 percent, $p < 0.01$, respectively). These reductions started in Year 3 and continued through Year 9. This is consistent with the slower growth in outpatient ED visits during the analysis period. It should be noted, though, that expenditures on outpatient ED visits were a small component of total outpatient expenditures.
- There was a 8.9 percent (\$2.6 PBPM, $p = 0.03$) relative increase in hospice expenditures over the nine years, driven by an 11 percent increase in the post-Classic period (\$3.7 PBPM, $p = 0.01$). The yearly estimates were statistically significant in Year 3 (\$2.4 PBPM, 10.5 percent, $p = 0.07$), Year 5 (\$3.4 PBPM, 12.3 percent, $p = 0.01$), Year 7 (\$3.3 PBPM, 9.0 percent, $p = 0.08$), Year 8 (\$4.3 PBPM, 10.7 percent, $p = 0.02$), and Year 9 (\$6.0 PBPM, 14.5 percent, $p < 0.01$).
- There were no discernable differences between intervention and comparison practices in expenditures on physician services, home health, skilled nursing facilities, or durable medical equipment (DME) over the course of the nine years. Some yearly estimates were statistically significant but there were no consistent patterns.

Figure 5.H.3. Estimated effects on acute, non-acute, and total inpatient expenditures for attributed Medicare fee-for-service beneficiaries for CPC Classic and comparison practices, by year



Source: Medicare claims data for October 2011 through December 2021.

Notes: The estimate of the effect, denoted by bars of separate colors for total, acute, and non-acute inpatient expenditures in the figure, is equal to the difference in mean outcomes between attributed Medicare FFS beneficiaries in the intervention and comparison group practices in any year since CPC Classic began minus the average difference between the two groups during the baseline period. Estimates are regression-adjusted for baseline beneficiary characteristics, baseline practice characteristics, and COVID-19-related controls.

*/**/*** Significantly different from zero at the 0.10/0.05/0.01 levels, two-tailed test.

COVID-19 = coronavirus disease 2019; CPC = Comprehensive Primary Care; FFS = fee-for-service; Y = intervention year.

Table 5.H.5. Regression-adjusted means and difference-in-differences estimates for expenditures by service categories among attributed Medicare fee-for-service beneficiaries for CPC Classic and comparison practices, annual and cumulative estimates

	Intervention mean	Comparison mean	Difference-in-differences estimate (SE)	Difference-in-differences estimate in percentage[a]	90 percent confidence interval	*p*-Value
Medicare expenditures (PBPM)						
Total inpatient expenditures						
Baseline	$205	$201	NA	NA	NA	NA
Y1	$315	$316	-$4.7 ($5.0)	-1.5%	(-$12.9, $3.5)	0.35
Y2	$321	$321	-$3.7 ($5.0)	-1.1%	(-$11.9, $4.6)	0.46
Y3	$330	$321	$4.9 ($4.7)	1.5%	(-$2.8, $12.7)	0.30
Y4	$320	$317	-$0.4 ($4.9)	-0.1%	(-$8.5, $7.6)	0.93
Y5	$321	$319	-$1.6 ($4.5)	-0.5%	(-$9.0, $5.9)	0.73
Y6	$330	$326	-$0.1 ($5.3)	0.0%	(-$8.9, $8.6)	0.98
Y7	$352	$347	$0.7 ($5.6)	0.2%	(-$8.5, $9.9)	0.90
Y8	$337	$339	-$6.2 ($5.7)	-1.8%	(-$15.6, $3.1)	0.27
Y9	$365	$360	$1.1 ($6.8)	0.3%	(-$10.1, $12.2)	0.88
Y1–Y4 (CPC Classic period)	$322	$319	-$0.7 ($3.9)	-0.2%	(-$7.2, $5.8)	0.86
Y5–Y9 (Post-Classic period)	$339	$336	-$1.2 ($4.5)	-0.4%	(-$8.6, $6.1)	0.78
Y1–Y9	$331	$328	-$1.0 ($4.0)	-0.3%	(-$7.5, $5.5)	0.80
Acute inpatient						
Baseline	$178	$178	NA	NA	NA	NA
Y1	$275	$280	-$5.3 ($4.5)	-1.9%	(-$12.7, $2.0)	0.23
Y2	$280	$286	-$6.4 ($4.5)	-2.2%	(-$13.9, $1.0)	0.16
Y3	$288	$285	$3.0 ($4.5)	1.0%	(-$4.4, $10.3)	0.51
Y4	$281	$282	-$1.2 ($4.5)	-0.4%	(-$8.5, $6.2)	0.80
Y5	$280	$283	-$3.7 ($4.1)	-1.3%	(-$10.4, $3.1)	0.37
Y6	$288	$290	-$2.3 ($4.7)	-0.8%	(-$10.0, $5.4)	0.62
Y7	$305	$310	-$4.3 ($5.0)	-1.4%	(-$12.6, $3.9)	0.39
Y8	$291	$301	-$9.2* ($4.9)	-3.1%	(-$17.3, -$1.1)	0.06
Y9	$315	$319	-$4.2 ($6.0)	-1.3%	(-$14.1, $5.7)	0.48
Y1-Y4 (CPC Classic period)	$281	$283	-$2.2 ($3.6)	-0.8%	(-$8.2, $3.7)	0.54

Table 5.H.5. ***(continued)***

	Intervention mean	Comparison mean	Difference-in-differences estimate (SE)	Difference-in-differences estimate in percentage[a]	90 percent confidence interval	*p*-Value
Y5-Y9 (Post-Classic period)	$294	$299	-$4.6 ($3.9)	-1.5%	(-$11.1, $1.9)	0.24
Y1-Y9	$288	$292	-$3.5 ($3.6)	-1.2%	(-$9.4, $2.4)	0.33
Non-acute inpatient						
Baseline	$27	$23	NA	NA	NA	NA
Y1	$40	$35	$0.7 ($1.5)	1.7%	($-1.8, $3.1)	0.66
Y2	$42	$35	$2.7* ($1.6)	7.0%	($0.1, $5.4)	0.09
Y3	$43	$37	$2.0 ($1.5)	4.8%	($-0.5, $4.4)	0.18
Y4	$40	$35	$0.7 ($1.3)	1.9%	($-1.4, $2.9)	0.57
Y5	$42	$36	$2.1 ($1.5)	5.2%	($-0.4, $4.5)	0.16
Y6	$42	$36	$2.2 ($1.6)	5.4%	($-0.4, $4.8)	0.17
Y7	$47	$38	$5.1*** ($1.8)	12.2%	($2.0, $8.1)	0.01
Y8	$46	$39	$2.9 ($1.9)	6.9%	($-0.2, $6.1)	0.13
Y9	$49	$40	$5.3** ($2.0)	12.0%	($1.9, $8.6)	0.01
Y1-Y4 (CPC Classic period)	$41	$36	$1.5 ($1.2)	3.9%	($-0.5, $3.5)	0.21
Y5-Y9 (Post-Classic period)	$45	$37	$3.4** ($1.4)	8.1%	($1.1, $5.6)	0.01
Y1-Y9	$43	$37	$2.5** ($1.2)	6.2%	($0.5, $4.5)	0.04
Outpatient						
Baseline	$97	$103	NA	NA	NA	NA
Y1	$116	$123	-$1.7 ($1.4)	-1.5%	(-$4.0,$0.6)	0.23
Y2	$128	$137	-$2.5 ($1.8)	-1.9%	(-$5.5,$0.4)	0.16
Y3	$138	$148	-$4.0** ($1.8)	-2.8%	(-$6.9,-$1.1)	0.02
Y4	$147	$156	-$3.7* ($2.0)	-2.5%	(-$7.0,-$0.4)	0.07
Y5	$162	$176	-$7.8*** ($2.6)	-4.6%	(-$12.1,-$3.6)	0.00
Y6	$178	$190	-$5.9** ($2.8)	-3.2%	(-$10.5,-$1.2)	0.04
Y7	$194	$208	-$8.5*** ($3.1)	-4.2%	(-$13.5,-$3.4)	0.01
Y8	$184	$196	-$6.6* ($3.6)	-3.5%	(-$12.5,-$0.6)	0.07
Y9	$216	$230	-$8.4* ($4.3)	-3.8%	(-$15.5,-$1.3)	0.05
Y1-Y4 (CPC Classic period)	$133	$142	-$3.1** ($1.4)	-2.3%	(-$5.4,-$0.7)	0.03
Y5-Y9 (Post-Classic period)	$185	$198	-$7.4*** ($2.6)	-3.9%	(-$11.7,-$3.1)	0.00

Table 5.H.5. ***(continued)***

	Intervention mean	Comparison mean	Difference-in-differences estimate (SE)	Difference-in-differences estimate in percentage[a]	90 percent confidence interval	*p*-Value
Y1-Y9	$161	$172	-$5.4*** ($1.9)	-3.2%	(-$8.5,-$2.3)	0.00
Physician						
Baseline	$200	$195	NA	NA	NA	NA
Y1	$228	$223	-$0.2 ($1.7)	-0.1%	(-$2.9,$2.6)	0.92
Y2	$233	$229	-$1.3 ($1.8)	-0.5%	(-$4.3,$1.8)	0.49
Y3	$243	$237	$1.6 ($2.0)	0.7%	(-$1.6,$4.9)	0.40
Y4	$252	$242	$4.7** ($2.4)	1.9%	($0.8,$8.6)	0.05
Y5	$258	$249	$4.7* ($2.7)	1.9%	($0.3,$9.2)	0.08
Y6	$268	$261	$2.5 ($3.1)	0.9%	(-$2.6,$7.6)	0.43
Y7	$292	$280	$6.9* ($3.6)	2.4%	($1.0,$12.8)	0.05
Y8	$275	$268	$2.0 ($3.5)	0.7%	(-$3.8,$7.7)	0.57
Y9	$326	$310	$12.0*** ($4.5)	3.8%	($4.7,$19.4)	0.01
Y1–Y4 (CPC Classic period)	$240	$234	$1.4 ($1.6)	0.6%	(-$1.2,$4.0)	0.37
Y5–Y9 (Post-Classic period)	$281	$271	$5.3* ($3.0)	1.9%	($0.3,$10.3)	0.08
Y1–Y9	$262	$254	$3.5 ($2.2)	1.4%	(-$0.1,$7.1)	0.11
Home health						
Baseline	$26	$30	NA	NA	NA	NA
Y1	$39	$44	-$1.3** ($0.6)	-3.3%	(-$2.4,-$0.3)	0.03
Y2	$40	$43	$0.8 ($0.7)	2.0%	(-$0.4,$2.0)	0.27
Y3	$42	$45	$0.3 ($0.7)	0.8%	(-$0.9,$1.6)	0.64
Y4	$41	$46	-$0.4 ($0.9)	-0.9%	(-$1.8,$1.1)	0.69
Y5	$43	$48	-$1.1 ($1.0)	-2.5%	(-$2.7,$0.4)	0.24
Y6	$46	$52	-$2.2** ($1.0)	-4.6%	(-$3.9,-$0.5)	0.03
Y7	$47	$54	-$2.7** ($1.0)	-5.4%	(-$4.4,-$1.0)	0.01
Y8	$44	$49	-$0.9 ($1.2)	-2.0%	(-$2.9,$1.1)	0.46
Y9	$50	$55	-$1.0 ($1.2)	-1.9%	(-$2.9,$1.0)	0.42
Y1–Y4 (CPC Classic period)	$41	$45	-$0.1 ($0.6)	-0.2%	(-$1.1,$0.9)	0.87
Y5–Y9 (Post-Classic period)	$46	$51	-$1.6* ($0.9)	-3.4%	(-$3.1,-$0.1)	0.07
Y1–Y9	$43	$48	-$0.9 ($0.7)	-2.0%	(-$2.0,$0.2)	0.20

Table 5.H.5. ***(continued)***

	Intervention mean	Comparison mean	Difference-in-differences estimate (SE)	Difference-in-differences estimate in percentage[a]	90 percent confidence interval	*p*-Value
Hospice						
Baseline	$2	$2	NA	NA	NA	NA
Y1	$20	$20	$0.4 ($1.0)	1.9%	(-$1.3,$2.0)	0.72
Y2	$23	$23	$0.4 ($1.3)	1.9%	(-$1.7,$2.5)	0.74
Y3	$25	$23	$2.4* ($1.3)	10.5%	($0.2,$4.6)	0.07
Y4	$27	$26	$2.0 ($1.3)	7.8%	(-$0.2,$4.1)	0.13
Y5	$31	$28	$3.4*** ($1.3)	12.3%	($1.3,$5.6)	0.01
Y6	$35	$34	$2.3 ($1.6)	7.0%	(-$0.3,$5.0)	0.15
Y7	$40	$38	$3.3* ($1.9)	9.0%	($0.2,$6.5)	0.08
Y8	$45	$41	$4.3** ($1.9)	10.7%	($1.2,$7.4)	0.02
Y9	$48	$42	$6.0*** ($1.9)	14.5%	($2.8,$9.2)	0.00
Y1–Y4 (CPC Classic period)	$24	$23	$1.4 ($1.1)	6.0%	(-$0.4,$3.2)	0.21
Y5–Y9 (Post-Classic period)	$39	$36	$3.7*** ($1.4)	10.6%	($1.4,$6.1)	0.01
Y1–Y9	$32	$30	$2.6** ($1.2)	8.9%	($0.6,$4.6)	0.03
Skilled nursing facility						
Baseline	$30	$32	NA	NA	NA	NA
Y1	$61	$68	-$4.6*** ($1.7)	-7.0%	(-$7.4,-$1.8)	0.01
Y2	$64	$70	-$4.1** ($1.8)	-6.0%	(-$7.0,-$1.2)	0.02
Y3	$68	$72	-$2.1 ($2.0)	-3.0%	(-$5.4,$1.1)	0.28
Y4	$66	$70	-$1.7 ($2.1)	-2.6%	(-$5.1,$1.6)	0.40
Y5	$68	$74	-$2.9 ($2.2)	-4.1%	(-$6.6,$0.7)	0.18
Y6	$71	$77	-$3.4 ($2.5)	-4.5%	(-$7.4,$0.7)	0.17
Y7	$76	$79	-$1.3 ($2.6)	-1.7%	(-$5.6,$3.0)	0.62
Y8	$81	$78	$4.8* ($2.8)	6.3%	($0.2,$9.3)	0.09
Y9	$86	$86	$1.9 ($2.8)	2.2%	(-$2.7,$6.5)	0.50
Y1–Y4 (CPC Classic period)	$65	$70	-$3.0* ($1.7)	-4.4%	(-$5.7,-$0.3)	0.07
Y5–Y9 (Post-Classic period)	$76	$78	-$0.5 ($2.1)	-0.7%	(-$4.0,$2.9)	0.79
Y1–Y9	$71	$75	-$1.7 ($1.8)	-2.4%	(-$4.7,$1.3)	0.35

Table 5.H.5. ***(continued)***

	Intervention mean	Comparison mean	Difference-in-differences estimate (SE)	Difference-in-differences estimate in percentage[a]	90 percent confidence interval	*p*-Value
Durable medical equipment						
Baseline	$23	$23	NA	NA	NA	NA
Y1	$25	$26	$0.1 ($0.4)	0.4%	(-$0.5,$0.7)	0.81
Y2	$22	$23	-$0.5 ($0.5)	-2.1%	(-$1.3,$0.4)	0.35
Y3	$23	$24	-$0.9* ($0.5)	-3.8%	(-$1.8,-$0.0)	0.09
Y4	$21	$23	-$1.0* ($0.6)	-4.5%	(-$2.0,-$0.1)	0.08
Y5	$21	$22	-$0.9 ($0.7)	-4.1%	(-$2.0,$0.2)	0.18
Y6	$23	$24	-$0.3 ($0.7)	-1.3%	(-$1.5,$0.9)	0.69
Y7	$25	$27	-$1.2 ($0.9)	-4.4%	(-$2.6,$0.3)	0.18
Y8	$27	$28	-$1.0 ($0.9)	-3.5%	(-$2.5,$0.6)	0.30
Y9	$27	$29	-$1.3 ($1.0)	-4.4%	(-$2.9,$0.4)	0.21
Y1–Y4 (CPC Classic period)	$22	$24	-$0.6 ($0.4)	-2.7%	(-$1.3,$0.1)	0.13
Y5–Y9 (Post-Classic period)	$24	$26	-$0.9 ($0.7)	-3.5%	(-$2.1,$0.3)	0.22
Y1–Y9	$23	$25	-$0.8 ($0.5)	-3.1%	(-$1.6,$0.1)	0.14
Sample sizes						
Number of practices	497	908				
Number of beneficiaries	565,674	1,165,284				
Number of beneficiary years	3,955,515	8,118,902				

Source: Medicare claims data for October 2011 through December 2021.

Notes: Estimates are regression adjusted for baseline beneficiary characteristics, baseline practice characteristics, and COVID-19-related controls. We based each estimate on a difference-in-differences analysis, and each reflects the difference in the regression-adjusted average outcome for attributed Medicare FFS beneficiaries in intervention practices in Years 1 to 7 compared with baseline relative to the same difference over time for attributed Medicare FFS beneficiaries in comparison practices. Note that expenditures are generally lower in the baseline year (relative to intervention years) because the baseline sample is composed of beneficiaries who were attributed (and hence alive) during the intervention period and did not include beneficiaries who needed expensive end-of-life care who would have died during the baseline year.

[a] To calculate these percentages, we divided the difference-in-differences estimate by the mean for the outcome in the intervention group minus the difference-in-differences estimate.

*/**/*** Significantly different from zero at the 0.10/0.05/0.01 levels, two-tailed test.

COVID-19 = coronavirus disease 2019; CPC = Comprehensive Primary Care; FFS = fee-for-service; n.a. = not applicable; PBPM = per beneficiary per month; SE = standard error; Y = intervention year.

5.H.4. Discussion

Results from this analysis provide the first estimates of longer-term effects of primary care transformation on expenditures and service use outcomes. We examined nine years of expenditures and utilization data, combining four years of CPC Classic, followed by five years of CPC+ for many practices, and have three main findings:

- The intervention reduced growth in acute hospitalizations over the full nine-year period by 2 percent. This was driven by a statistically significant relative reduction in the post-Classic period (3 percent).
- In addition, the intervention reduced growth in outpatient ED visits (approximately 2 percent) both during the CPC Classic model period and after CPC Classic ended.
- There was no discernible difference between the intervention and comparison practices in Medicare Part A and B expenditures excluding enhanced.

The temporal pattern of effects on outpatient ED visits and acute hospitalizations is consistent with our expectations about how primary care transformation works—outcomes like outpatient ED visits could be easier to improve in the short run, which would explain the emergence of favorable effects during the four-year CPC Classic period, whereas a longer time horizon may be needed to see improvements in outcomes like acute hospitalizations. Because many CPC Classic practices (85 percent) joined CPC+ in 2017 and many of their comparison practices (79 percent) did not join CPC+ in 2017, these favorable effects reflect the combined effects of the four years of CPC Classic and the five years of CPC+. We cannot determine how much of the effects in the post-Classic period are attributable to the lagged effects of CPC Classic versus the additional years of support through CPC+. It is likely that CPC+ provided important support to continue the work begun in CPC Classic for the CPC Classic practices that joined CPC+.

Although the relative reductions in acute hospitalizations and outpatient ED visits are promising, they did not translate to a discernable relative reduction in overall expenditures. There are two potential explanations:

There were relative increases in hospice expenditures, physician expenditures, and non-acute inpatient expenditures that offset the relative reductions in acute inpatient expenditures and outpatient expenditures.

The increase in hospice expenditures could be explained by the requirements of CPC Classic and CPC+ for practices to improve end-of-life planning. As for the increase in non-acute inpatient expenditures, there could be potential substitutions between acute and non-acute inpatient care, which in some cases might be beneficial to patients' health. For example, CPC's emphasis on care coordination and care management, which included transitional care planning, may increase the appropriate recommendation of inpatient rehabilitation stays (which is part of non-acute inpatient care). This could prevent premature return to the home environment, ensure appropriate home modifications as well as patient's greater physical resilience once back at home, and thus avoid future acute hospitalizations. Consistent with this, in the impact analyses of the CPC+ final report, we found an increase in inpatient rehabilitation expenditures across the five program years (5 percent for Track 1, $p < 0.01$; 7.5 percent for Track 2, $p < 0.01$, see Appendix 5.A of the CPC+ final report).

Compared to the relative reduction in acute hospitalizations (3 percent) during the post-Classic period, the estimated effect on acute inpatient expenditures was smaller (1.5 percent) and not statistically significant ($p = 0.24$). In the impact analyses of the CPC+ final report, we examined the effect of CPC+ on different types of acute hospitalizations and found that the reductions in acute hospitalizations and inpatient expenditures were driven by reductions in medical (non-surgical) admissions and admissions without any complication or comorbidity (see Appendix 5.J of the CPC+ final report). Although we have not examined the effects of the CPC models on types of acute hospitalizations specifically for this analysis, the findings from the CPC+ impact analyses provide supportive evidence for the hypothesis that the avoided hospitalizations could be relatively less severe and less costly, thus explaining the relatively smaller reduction in acute inpatient expenditures.

Even the effects on acute hospitalizations and outpatient ED use that we do observe are modest in size. It is possible that effects might be larger if primary care practices had stronger incentives or if there were incentives for other providers (for example, including hospitals and specialists) who care for the same patients. Also, beneficiaries were not rewarded for taking better care of themselves or seeking higher-value providers or services. Finally, comparison practices' outcomes may have improved due to other efforts to transform primary care (for example, through the increase in penalties for high readmission rates); this may have made it difficult for the intervention practices to generate reductions in savings or service use relative to the comparison practices.

This study has several limitations. First, as with any difference-in-differences study design, there are concerns about the validity of the underlying assumption of parallel trends in outcomes. One specific concern related to our approach to matching is that we matched on outcomes during the pre-intervention period to select our comparison practices. In a study published in 2018 (two years after the end of CPC Classic), Daw and Hatfield showed that regression to the mean can lead to bias in studies with comparison group designs that match on pre-intervention outcomes. However, they also pointed out that this issue is most problematic when the difference in outcomes between potential comparisons and selected comparisons is large. In this analysis, the difference in average outcome values in the group of potential comparison practices (pre-matching) and selected comparison practices (post-matching) were small, suggesting that regression to the mean is not likely to substantially bias our results (Dale et al. 2016). Another concern is due to the long evaluation period (nine years) used in this study. With such a long study period, even small differences in the outcomes trends at baseline, if they persisted linearly at the same rate, could potentially compound over time and confound the impacts of the intervention. While we have some evidence that mitigates this concern, we cannot rule it out completely. Specifically, Fu et al. (2022) showed there was no noticeable trend in the difference in outcomes between the intervention and comparison groups over the two years before CPC Classic began. In addition, Dale et al. (2016) showed that impact estimates during the CPC Classic period were robust to changes in the length of the baseline period (from one to two years), which likely would have shown different results if baseline differences in trends were driving impacts. It should also be noted that the impacts on hospitalizations that drive our key takeaway emerged in Year 5, roughly halfway through the intervention period (and not towards the end of the nine-year study period). Further, the annual estimated effects in Years 5 through 7 (3.2 percent to 4.1 percent, $p < 0.01$) were distinctly different from the estimates in Years 1 through 4 (up to 2 percent [Year 2], not statistically significant at the 10 percent level), which suggests that the steady compounding of baseline differences is not driving our key finding.

Second, we should interpret our estimates for 2020 and 2021 with caution due to the unforeseen COVID-19 pandemic. CPC Classic regions experienced a 1.4 percent larger reduction in Medicare Part A and Part B expenditures from 2019 to 2020 compared to CPC Classic external comparison regions, indicating that COVID-19 might have affected CPC Classic and comparison practices differently. Although we controlled for COVID-19-related regional factors in a flexible and granular way (consistent with the impact analyses in the CPC+ final report), we did not assess the appropriateness of the COVID-19-related controls specifically for this analysis and there remains uncertainty about potential bias caused by unobserved COVID-19-related factors that we would not be able to assess.

Third, due to our intent-to-treat design, some beneficiaries assigned to the intervention group no longer receive care at practices participating in CPC Classic in later follow-up years, potentially leading to attenuation bias. In addition, 21 percent of CPC Classic comparison practices joined CPC+ and although the beneficiaries assigned to these practices potentially benefited from CPC+, they remained in the comparison group in Years 5 through 9, also potentially leading to an underestimate of the full extent of the combined intervention's favorable effect on outcomes. Given the long evaluation period (nine years) and the modest estimated impact size, it is possible that some true impacts were not detected because of these attenuation biases.

Finally, findings from the CPC models, with the unique set of practices and patients, may not generalize to other payers, primary care models, or participants with different eligibility requirements, model rules, and supports.

5.H.5. Conclusion

The findings from this analysis have important implications for how payers and policymakers should test and assess primary care reform over longer periods. The results suggest that primary care transformation may reduce outpatient ED visits quickly, that it could take a longer period of robust support to reduce acute hospitalizations, and that reducing total health care spending may require longer or new approaches. More research is needed to assess longer-term effects of other primary care transformation approaches to see if similar temporal patterns appear.

5.I. Scalability

Key takeaways

In this Appendix, we project the impact CPC+ could have if CMS were to scale the model. Specifically, we generalize the effect of the last year of the five-year model, accounting for CPC+ practices potentially not being representative of those in a scale-up. We consider both a *nationwide scale-up* and a *targeted scale-up* to practices where the intervention would be most likely to generate savings. We estimate these scale-up impacts using weighted Bayesian Causal Forests (wBCF). Through its data-driven discovery of impact heterogeneity, wBCF can identify the populations expected to benefit most from an intervention, informing decisions about targeted scale-up. Our key findings are as follows:

- There is almost no chance a nationwide scale-up of either track's fifth-year effects would generate sufficient savings to offset care management fees (CMFs).
- We likewise did not find any targeted scale-up likely to be cost neutral. A Track 1 scale-up to Medicare Shared Savings Program (SSP) practices—which showed the most promise in Program Year (PY) 4, with a 79 percent probability of offsetting CMFs—likewise showed the most favorable effects in PY 5. However, these PY 5 effects were attenuated and showed only a 34 percent probability of a cost-neutral scale-up.
- A targeted scale-up of Track 1 to SSP practices would also likely decrease both outpatient ED visits (86 percent probability) and acute hospitalizations (89 percent probability).

Although our analysis extrapolated CPC+ impacts geographically, it did not tackle the more difficult challenge of extrapolating impacts forward in time: we assessed the impact of CPC+ *as it was offered in PY 5.* This appendix details methods, their assumptions, full results, sensitivity analyses, and limitations of the scalability analysis.

5.I.1. Introduction

In this Appendix, we estimate the impact if CMS were to scale up CPC+, focusing on three key outcomes: (1) total Medicare expenditures without enhanced CPC+ payments, (2) outpatient emergency department (ED) visits including observation stays, and (3) acute hospitalization rates. Although Chapter 5 estimates the model's impacts on expenditures for the practices that joined in 2017, scaling up to a set of practices with a different profile of practice and patient characteristics might show different effects. We consider both a *nationwide scale-up* to all practices that would be eligible in the United States across both CPC+ regions and new regions and a *targeted scale-up* to practices in which the intervention would be most likely to generate savings.

The primary CPC+ impact analysis provides reliable estimates of CPC+'s effects in the *evaluation sample*—that is, among the set of practices that began participating in CPC+ in 2017. CMS selected regions and then practices in those regions that were motivated to apply to the model and met CMS's eligibility criteria. However, this set of practices might not be representative of the eligible practices that would volunteer for a scale-up—that is, the *projected scaled sample* that would represent the target population.

Evaluation sample: practices that participated in Track 1 or Track 2 of CPC+

Projected scaled sample: practices that would be eligible and would volunteer to participate in a scale-up

As a result, we cannot interpret the evaluation sample's impact estimate (presented in Chapter 5) as the impact we would expect if CPC+ were scaled up. To gauge the effects of CPC+ in this larger sample of practices, we must account for differences in practice and patient characteristics between the evaluation sample and the projected scaled samples—namely, differences in characteristics that modify the effect of CPC+. For example, if CPC+ generated the largest savings among practices participating in the Medicare Shared Savings Program (SSP) and if the proportion of practices that participate in SSP differs between the evaluation sample and the projected scaled sample, we would expect average CPC+ impacts under a scale-up to differ from those in the evaluation. Thus, even when the impact evaluation has dealt *with internal validity biases* (confounding and other factors that could bias in-sample estimates of the effects of CPC+ for the evaluation sample), to recover the true impacts in the scale-up, we still need to address *external validity biases* (differences in practice and beneficiary characteristics that could bias out-of-sample estimates of the effects of CPC+).

We cannot interpret the evaluation sample's impact estimate as the impact we would expect if CPC+ were scaled up, because the evaluation sample and projected scaled sample might differ on characteristics that modify the effect of CPC+.

To address the discrepancy between the evaluation sample and projected scaled samples and extend impact results beyond the evaluation at hand, we must use generalizability methods. These methods have attracted increasing attention, resulting in approaches that use outcome regressions such as ordinary least squares models or Bayesian additive regression trees (BART; Hill 2011; Green and Kern 2012; Kern et al. 2016); propensity of selection weighting approaches (Cole and Stuart 2010; Correa et al. 2018); and double-robust estimators, such as the targeted maximum likelihood estimator (Rudolph and van der Laan 2017) and augmented inverse probability weighting (Dahabreh et al. 2018). However, other than BART, most approaches have relied on parametric modeling assumptions that relationships between covariates and the outcome are linear and additive. Few existing approaches allow for flexible modeling, which is particularly important when generalizing results from large observational studies with many confounders and effect modifiers, studies in which the true relationships between covariates and the outcome are unknown and not easily captured through simple linear additive relationships.

Estimating the impact of a CPC+ scale-up requires additional considerations novel to the generalizability literature because of the voluntary nature of model participation. Namely, because scale-up participation would remain voluntary, the projected scaled samples are uncertain: it is unclear which practices would be eligible and would volunteer for a scale-up. Although there is a large body of literature on considerations for scaling up implementation (Barker et al. 2016; Powell et al. 2015; World Health Organization 2010), and several approaches exist for estimating impacts of a policy model scale-up (Attanasio et al. 2003; Flores and Mitnik 2013; Gechter 2015), to our knowledge, no literature addresses generalizability to a projected scaled sample that is not enumerable (because of uncertainty about which practices would volunteer for the scaled-up intervention in new geographic regions).

To address these shortcomings, we present a novel approach to scalability analysis that uses nonparametric outcome regressions called Bayesian Causal Forests (BCF) to extend inference from the CPC+ evaluation sample to the projected scaled sample, while accounting for effect heterogeneity across different types of practices in a data-driven fashion. BCF has shown superior performance over other causal estimators for confounding adjustment (Hahn et al. 2019). Also, BCF extends particularly naturally to estimating scale-up effects as it explicitly considers and accounts for both confounding (the key threat to internal validity) and heterogeneous treatment effects (the key threat to external validity). Specifically,

we estimated impacts of scaling CPC+, using Program Year (PY) 5 outcomes, with a version of BCF called weighted BCF (wBCF), which allows larger practices to contribute more information than smaller practices. We focused on PY 5 outcomes because we believed PY 5 effects would represent the cumulative effect of the care delivery changes made and sustained during the model implementation. However, we recognize estimated effects in PY 5 could be highly sensitive to variation because of coronavirus 2019 (COVID-19).

As Bayesian estimators, BCF methods (including wBCF) allow for incorporating additional sources of uncertainty into the credible intervals, such as uncertainty about what factors will drive participation in new geographic regions. Incorporating these sources of uncertainty into the final Bayesian credible intervals will help us avoid overstating confidence in estimated scale-up effects. Thus, our point estimates and uncertainty bounds directly capture uncertainty about which practices will volunteer for a scale-up. We address additional sources of uncertainty in sensitivity analyses, namely uncertainty regarding unmeasured effect modifiers such as motivation to improve care, assumptions made by different modeling choices (BCF versus the difference-in-differences [DD] regression used in the main impact analysis in Chapter 5), and the potential for the COVID-19 pandemic to confound our PY 5 estimates. In this analysis, we assessed the long-term impact of scaling up Tracks 1 and 2 of CPC+ *as they were offered during the evaluation,* through the final program year, PY 5. This approach limits the extent to which our scale-up estimates can *predict* what would happen in a future scale-up, because CPC+ implementation would undoubtedly change under scale-up. Instead, our estimates, though attempting to forecast the future, can be more precisely thought of as providing an accurate *retrospective* estimate of the impact if CPC+ been offered nationwide in 2017. We consider this limitation in more detail in the Discussion section of this Appendix.

Furthermore, a practical challenge of implementing targeted scale-up approaches is that practice baseline characteristics have changed since they were assessed before the evaluation began; for example, the set of practices participating in SSP today differs from the set that participated in 2017. Our analysis does not incorporate uncertainty about extrapolating *practice characteristics* forward in time. The Discussion section considers this and additional limitations.

To obtain scale-up estimates, the analysis proceeded in four steps:

1. For the nationwide and targeted scale-ups, we estimated which eligible practices would participate.
2. We used wBCF to estimate the PY 5 impact of CPC+ in fine-grained subgroups within the evaluation sample of CPC+ practices and their matched comparison practices, separately for Track 1 and Track 2.
3. Using the wBCF fitted regression, we predicted CPC+'s impact in the projected scaled samples.
4. We assessed sensitivity of these predicted scaled impact estimates to assumptions.

5.I.2. Methods

The scalability analysis addressed what would have happened if, at the end of the evaluation period, CMS had scaled up CPC+ instead of discontinuing the model. The outcomes of interest were total Medicare expenditures without enhanced CPC+ payments, outpatient ED visits including observation stays, and acute hospitalization rates. We examined the causal impact of CPC+ as it was implemented in practices that joined CPC+ in 2017. Specifically, we estimated the effects of both a *nationwide scale-up* and a *targeted scale-up* to practices in which the intervention would be likely to generate savings. The

nationwide scale-up explored scaling up to all eligible practices nationwide that would volunteer, across both CPC+ regions and new regions. Practice-targeting scale-ups explored scaling up to practice types in which wBCF estimated total Medicare expenditure savings from CPC+ to be most likely to offset care management fees (CMFs).

Step 1. Identify CPC+-eligible practices nationwide that would participate

Within CPC+ regions, practices could volunteer for one of two model tracks. The model enrolled all volunteering practices that met eligibility requirements. Because a nationwide scale-up would remain voluntary for practices that meet eligibility requirements, we had to identify practices that would be eligible for CPC+ and were likely to volunteer to participate. These eligible volunteering practices—either nationwide or among targeted practices—represent the projected scaled samples to which we wish to generalize impacts of the CPC+ model.

We started with the same nationwide data set of primary care practices used for comparison group selection; Singh et al. 2020 and Appendix 6.C in the second annual report (Ghosh et al. 2020) contain more details on how this data set was constructed. A small portion of practices (3.2 percent) were missing values for at least one covariate; we imputed these values from data on all primary care practices.

A. Determining eligibility

To determine which primary care practices in the United States would be eligible for a CPC+ scale-up, we approximated the claims-based eligibility criteria but not CPC+ care delivery requirements, as CMS determined care delivery criteria based on practices' applications, which were not available for practices nationwide (Table 5.I.1). The main impact analysis also approximated application-based criteria using the Medicare claims and other administrative data to narrow the set of potential comparison practices before matching (Peikes et al. 2021b).We deemed practices eligible if they provided primary care, at least 9.25 percent of their billed charges were for primary care services, they were not federally qualified health centers or rural health clinics, they were not participating in a Next Generation Accountable Care Organization (ACO) model as of January 2017, and they had more than 50 assigned Medicare fee-for-service (FFS) beneficiaries at the start of CPC+ (for stability of aggregate patient characteristics). These criteria differ from eligibility requirements CMS used because of the need to rely on data sources available for all practices nationwide to construct practice characteristics and attribute beneficiaries to practices; we could not rely on the data CMS used to assess eligibility and conduct attribution, because those data were available only for CPC+ practices (Table 5.I.1). Because we did not account for care delivery requirements, we likely considered more practices nationwide eligible than would truly meet all CPC+ eligibility requirements (e.g., 19 percent of CPC+ applicants did not meet one or more care delivery requirements). We did not apply the eligibility criterion of sufficient revenue from Medicare and other CPC+ payers, as CMS might not consider restricting practices based on payer alignment. Furthermore, because CMS provided 93 percent of unique funding for CPC+ practices in PY 5, payer alignment is likely to have had minimal impact on the model.

Table 5.I.1. Data sources to determine CPC+ eligibility in national sample

CMS eligibility criterion	Evaluation's modified criterion	Data source
Criteria we could approximate		
Primary care practice		SK&A 2016[a]
At least 40% of Medicare FFS services billed by the primary care practitioners are for primary care	≥ 9.25 primary care billing percentage[b]	Medicare claims data 2016
Not an FQHC or RHC		Provider of Service file 2016[c]
Not in Medicare shared savings aside from SSP (not in Next Generation ACO Model) on January 1, 2017		MDM 2017[d]
At least 125 Medicare beneficiaries attributed	> 50 Medicare beneficiaries assigned[e]	Medicare enrollment data 2015–2017; Medicare claims data 2014–2016
Uses CEHRT		Assumed 100%[f]
Criteria we could not approximate		
At least 45% of revenue comes from Medicare and CPC+ payer partners		Will no longer be an eligibility criterion
Patients assigned to provider panel		Unavailable
Patients have 24/7 access to a care team practitioner		Unavailable
Nonphysician team members deliver clinical care		Unavailable
Quality improvement activities		Unavailable

[a] We removed practices from the pool that we considered ineligible for CPC+ due to their intended patient populations. Specifically, we manually removed all practices that appeared to be specialty clinics (for example, surgery clinics, Planned Parenthood clinics, or urgent or emergency care clinics). We also removed practices with a practice specialty other than primary care, limiting the sample to the following eight specialties: (1) adolescent medicine, (2) family medicine, (3) geriatric medicine, (4) general practice, (5) internal medicine and pediatrics, (6) internal medicine, (7) multispecialty, and (8) pediatrics. (Pediatricians are not considered primary care physicians for CPC+. However, some practices with pediatric specialties participate in CPC+, because they have at least one practitioner with a primary care specialty; therefore, we included practices with pediatric or other specialties in our potential comparison sample as long as they had at least one practitioner with a nonpediatric primary care specialty.)

[b] The minimum billing percentage observed for CPC+ practices was 9.25 percent; under the methods used by the evaluation, some CPC+ practices had lower billing percentages than the eligibility threshold, because the evaluation computed the billing percentage using different data sources and a slightly different attribution algorithm than CMS.

[c] We did not assess FQHC and RHC status for all practices. Rather, at the time of analysis, we assessed this status only for practices selected to be comparison practices in an initial round of matching. FQHCs and RHCs could therefore remain in noncomparison practice regions. However, the minimum attributed beneficiary requirement makes it unlikely there were many.

[d] Participation status in Next Generation ACO model was missing for some practices (less than 1 percent of practices); we assumed these practices met eligibility requirements.

[e] We did not require at least 125 attributed beneficiaries because of differences between the attribution algorithms and data sources we used for the evaluation compared with those CMS uses for payment. Instead, we required more than 50 assigned beneficiaries for stability of aggregate patient characteristics, such as hospitalizations and Medicare expenditures.

[f] Due to the increase in EHR use over the past decade, including a requirement that all practices billing to Medicare use CEHRT, we conservatively assumed all practices would meet the CEHRT criterion at the end of the evaluation period.

ACO = Accountable Care Organization; CEHRT = Certified Electronic Health Record Technology; CMS = Centers for Medicare & Medicaid Services; EHR = electronic health record; FQHC = federally qualified health center; FFS = fee-for-service; MDM = master data management system; RHC = rural health clinic; SSP = Medicare Shared Savings Program.

B. Determining propensities to volunteer

CPC+ is a voluntary model: 15 percent of practices across all 18 CPC+ regions participated in the model. These participating practices were not representative of broader primary care practices in their region. For example, CPC+ practices were more likely to have patient-centered medical home recognition, participate in SSP, have meaningful use of an electronic health record (EHR), be owned by a system or hospital, and be larger than all practices providing primary care in their regions. The beneficiaries they served were slightly healthier and less disadvantaged than those whom all primary care practices in the CPC+ regions served (Singh et al. 2020).

Similarly, practices in an expanded voluntary model will not be representative of practices nationwide. To estimate which practices would volunteer to participate nationwide, we identified characteristics that drove eligible practices to participate in the 14 geographic regions that had already implemented CPC+ in 2017.[117] We did so by fitting a propensity score model for volunteering to participate in Track 1 or 2 among eligible practices in the 14 geographic regions, using multinomial BART. Then, we used the fitted model to predict the propensity that each eligible practice nationwide—that is, across both CPC+ regions and non-CPC+ regions—would volunteer for a scale-up of Track 1 or 2. Thus, we assumed CPC+ practices in the model would be highly likely but would not necessarily volunteer for the scale-up. We used the predicted propensities to participate as weights in subsequent steps of this analysis, such that practices nationwide with a higher predicted propensity to participate in a scale-up received more weight in our estimate of the scaled-up impact. We incorporated uncertainty from estimating the weights into the uncertainty intervals around our estimate of the scaled-up impact.

C. Targeting practices

Instead of scaling up CPC+ nationwide, CMS could restrict the scale-up to types of practices in which CPC+ would be most likely to generate savings. Therefore, in addition to estimating impacts of a Track 1 and Track 2 nationwide scale-up, for each track, we also estimated scale-up impacts for the subgroup of practices in which favorable expenditure impact estimates are most highly concentrated.

To estimate impacts of targeted scale-ups, we first assessed which practice characteristics best distinguish high-impact practices from low-impact practices using classification and regression trees (CART); these impacts were estimated in Steps 2 and 3. Specifically, we identified practice subgroups most likely to offset CMFs, which are $15 per beneficiary per month (PBPM) for Track 1 and $28 PBPM for Track 2. We focused on subgroups that improve CART's ability to explain impacts by at least 10 percent; when no subgroup met that criterion, we chose the subgroup most strongly explaining impacts. Then, if continuous variables defined high-impact practices, we determined thresholds for those characteristics to use as eligibility criteria for the practice-targeted scale-up, such that estimated savings would be large enough to offset CMFs. We then estimated expenditure, outpatient ED visit, and acute hospitalization scale-up impacts for these targeted practice subgroups.

Step 2. Estimate impact of CPC+ in fine-grained subgroups in the evaluation sample

For each track, we estimated impacts for two candidate projected scaled samples, corresponding to a nationwide and a practice-targeted scale-up. To do so, we first fitted wBCF regressions to the evaluation

[117] In this appendix, we do not analyze or report on the practices that joined CPC+ in 2018 for consistency with the practices used to estimate impacts, and as 2018 Starters account for only 5 percent of the total number of practices participating in CPC+.

sample (CPC+ practices) and their matched comparison practices in Tracks 1 and 2 to estimate CPC+ effects across different types of practices. Specifically, we fitted practice-level models (rather than a beneficiary-level models, to ensure computational tractability) that included the practice's average pre-period outcome as an independent variable (rather than as a dependent variable, as in the frequentist DD regressions presented in Chapter 5). The regression was similar to that used for the BCF scalability analysis in the fourth annual report (Appendix 5.I to Laird et al. 2022), with three key differences:

1. We used wBCF instead of the fourth annual report's individualized weighted BCF (iBCF) because for large subgroups such as the ones of interest for scalability analysis, wBCF's performance, including its uncertainty interval coverage, match those of iBCF across a range of settings. Given that performance is on par, we selected the simpler and better-established wBCF approach over the more complicated and less-established iBCF approach. In a sensitivity analysis, we found that PY 5 impact estimates were similar when using wBCF versus iBCF (results not shown).
2. We updated the region-level control variables to account for potential confounding and impact heterogeneity due to COVID-19 in 2021. Namely, we adjusted for the same covariates used in the main DD impact analysis to adjust for COVID-19 in PY 5 (see Appendix 5.D for further details).
3. To provide insight into the long-term effects of a scale-up, we analyzed PY 5 outcomes. To assess trends, we compared scalability findings using PY 5 estimates to findings using PY 3 and PY 4 estimates. We focus on PY 5 rather than cumulative impacts because we expect many effects to emerge in later years of the intervention, based on estimates of the longer-term effects of CPC Classic (Appendix 5.H) and previous literature on primary care transformation. However, COVID-19 might have disrupted this transformation, as the Discussion section explores. Furthermore, as BCF has not been extended to longitudinal analysis, a cumulative BCF analysis would have necessarily been fitted as a cross-sectional model of averages from PY 1 to PY 5; it therefore could not have appropriately adjusted for COVID-19, which affected PY 4 and PY 5 individually.

Step 3. Use fitted wBCF model to estimate CPC+ impact in projected scaled samples

In Step 3, we estimated causal effects of each track of CPC+ in each of the two projected scaled samples. We used the wBCF outcome regressions fit on the evaluation sample to compute a treatment effect estimate for each eligible practice in the projected scaled samples. The beneficiary-level population average treatment effects in each projected scaled sample are the propensity-for-volunteering and practice-size weighted averages of these estimated treatment effects. The propensity-for-volunteering weights are the predicted propensities for volunteering estimated in Step 1, and the practice-size weights enable larger practices to contribute more to the final estimates. Our approach is similar in spirit to weighted outcome regressions (Flores and Mitnik 2013) or multilevel regression with post-stratification (Gelman and Little 1997), though no existing approach accounts for estimating impacts of a voluntary intervention.

Step 4. Assess sensitivity to assumptions

In the final step, we assessed the robustness of wBCF's impact estimates to key external validity assumptions—those for projecting impacts to nationwide practices—and key internal validity assumptions—those for confounding adjustment. Namely, we assessed sensitivity of projected scaled sample estimates to (A) unmeasured effect modifiers. We assessed sensitivity of evaluation sample estimates to (B) using wBCF versus DD for confounding adjustment, and (C) to including an expanded set of comparison practices to better adjust for confounding by COVID-19 and other regional factors. Table 5.I.2 describes the sensitivity analyses we performed to address these three key sources of scale-up uncertainty.

Table 5.I.2. Sensitivity analyses

Sensitivity test	Motivation
A. Unmeasured effect modification	
Instead of assuming "no unmeasured effect modifiers," we assessed unmeasured effect modification as strong as (1) the strongest measured effect modifier bias, a worst-case scenario; and (2) the second-strongest effect modifier bias, a next-to-worst-case scenario.	This scalability analysis relies on canonical untestable assumptions related to internal and external validity (Degtiar and Rose 2023),* of which the *external validity* assumptions are novel to the scalability analysis (not shared by the main DD analysis). We expect this analysis to be most vulnerable to the assumption of *no unmeasured effect modification,* namely, that administrative data and secondary data used in the evaluation have captured all relevant factors affecting CPC+ impacts and that differ between evaluation sample and projected scale sample practices. However, the projected scaled samples could meaningfully differ from the evaluation sample in factors such as practices' motivation to improve patient care, baseline approaches to primary care delivery, and the payer landscape. We tested sensitivity of findings to such unmeasured effect modifiers.
B. Analytic approach comparison	
We compared evaluation sample impact estimates from wBCF with those from DD.	wBCF is a novel approach for estimating causal impacts and therefore not as time-tested as DD. To assess the sensitivity of findings to different confounding adjustment approaches, we compared evaluation sample impact estimates from wBCF with those from the frequentist DD presented in Chapter 5.
C. Expanded comparison group	
We expanded the comparison group to include all practices within CPC+ regions that were not enrolled in CPC+ and all practices within comparison regions (not just those selected for the comparison group used for the DD analysis).	Although BCF's flexibility enables it to control more fully than DD for confounding by COVID-19, this flexibility can also create collinearity (near-perfect alignment) between the regionally varying COVID-19 control variables and the regional variation in treatment status. Such collinearity can increase the width of BCF's credible intervals. This sensitivity analysis expanded the comparison group to all nonparticipating practices from CPC+ and comparison regions, lessening the collinearity between treatment status and regionally defined COVID-19 control variables, which should lead to more precise impact estimates. However, by adding unmatched practices to the comparison group, we more heavily rely on BCF's ability to appropriately adjust for confounding through its flexible modeling. The comparison practices used in this analysis parallel those used in the triple-differences (DDD) analysis presented in Appendix 5A. We compared evaluation sample estimates from this sensitivity analysis with the primary wBCF estimates and against DDD estimates.

*The full list of internal validity, external validity, and projected scaled samples assumptions and their implications for this analysis are as follows:

Internal validity assumptions

No unmeasured confounding with respect to treatment assignment. The CPC+ evaluation was not subject to unmeasured confounding: we adjusted for all variables that risk inducing internal validity bias if not appropriately accounted for.

Positivity of treatment assignment. All practices in the CPC+ evaluation had a positive probability of being in the CPC+ model.

Stable unit treatment value assumption (SUTVA) for treatment assignment. Practices did not affect each other's outcomes. Hence, there is no benefit or detriment from being in the same region as an existing CPC+ participant; although practices might have individually made different changes and received different payments as a result of participating in CPC+, the CPC+ model is well defined for all practices.

External validity assumptions

SUTVA for evaluation sample selection. Potential outcomes are not a function of how many practices are in the CPC+ model (no general equilibrium effects), practices will adopt the two tracks in similar ratios as we saw in the evaluation sample, the scale-up CPC+ model will not differ from the evaluation model (implementation by practices will remain the same), the same outcome relationships with covariates will hold in the scale-up, evaluation study participants will see annual benefits similar to those observed to date, and if CMS terminates the CPC+ model, these participants would revert to comparison group outcomes.

Table 5.I.2. ***(continued)***

No unmeasured confounding with respect to evaluation sample selection. There are no unmeasured effect modifiers related to being in one of the 14 geographic regions that implemented CPC+. Thus, we can expect new enrollees in the scale-up to benefit to a similar degree as evaluation CPC+ participants with similar measured characteristics. Sensitivity Analysis A assesses robustness of our estimates to violations of this assumption.

Positivity of sample selection. Eligible volunteering practices nationwide could have been in the CPC+ evaluation had it been implemented in their geography.

Projected scaled samples assumptions

Equivalent drivers of volunteering. Measured characteristics will determine which practices will volunteer for the scale-up in a similar way as they determined practice participation in CPC+ regions.

Invariant practice characteristics. Practice characteristics have not changed from the baseline period (2016–2017).

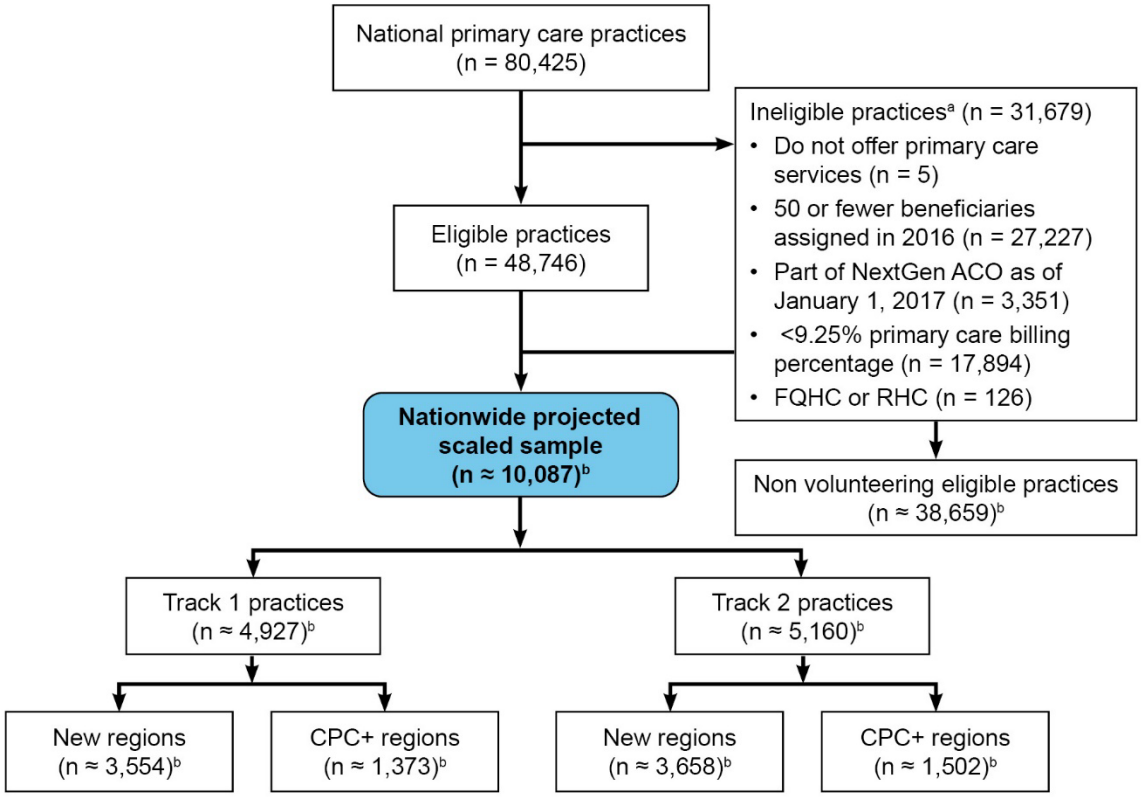
BCF = Bayesian Causal Forest; CMF = care management fee; CMS = Centers for Medicare & Medicaid Services; COVID-19 = coronavirus 2019; CPC+ = Comprehensive Primary Care Plus; DD = difference-in-differences; DDD = triple-differences; FFS = fee-for-service; PY = Program Year; wBCF = weighted Bayesian Causal Forest.

5.I.3. Results

Step 1. Identify CPC+-eligible practices nationwide that would participate

Of the 80,425 national primary care practices at the start of CPC+, we estimated 48,746 (61 percent) would be eligible to participate and, of eligible practices, about 4,927 (10 percent) would volunteer for Track 1 and 5,160 (11 percent) for Track 2 nationally (Figure 5.I.1). A similar proportion of practices volunteered for Tracks 1 and 2 in CPC+ regions (12 percent and 13 percent, respectively).

Figure 5.I.1. Number of primary care practices in the United States, by eligibility and projected decision to volunteer for a nationwide CPC+ scale-up



[a] The sample size for each reason for ineligibility includes practices excluded for multiple reasons.

[b] Approximate sample sizes are based on mean propensities to volunteer for each track.

ACO = Accountable Care Organization; CPC+ = Comprehensive Primary Care Plus; FQHC = federally qualified health center; n = number; RHC = rural health clinic.

Table 5.I.3 details the key differences between the evaluation sample, the projected scaled sample for a nationwide expansion of CPC+, and the remaining eligible practices in the United States we estimate would not volunteer for a nationwide scale-up.

Specifically, the practices that would be eligible and would volunteer for a nationwide scale-up of Track 1 differ as follows:

- **Compared with non-volunteering eligible practices.** We estimated sizable differences between the two samples across a number of characteristics. The nationwide projected scaled sample would have more assigned Medicare FFS beneficiaries on average compared with non-volunteering eligible practices (667 versus 417) and would be more likely to be owned or managed by a health system or

hospital (53 versus 30 percent), more likely to have attested to meaningful use of an EHR earlier (71 versus 49 percent attesting in 2011–2012), and more likely to have primary care transformation experience (35 versus 18 percent). We also estimated the nationwide projected scaled sample would have lower baseline expenditures ($898 versus $979), fewer acute care stays (283 versus 309 annually per 1,000 beneficiaries), and fewer ED visits (512 versus 557 annually per 1,000 beneficiaries). These practices would be more likely to have participated in an SSP than non-volunteering practices (55 versus 33 percent). Notably, this latter factor modifies the effect of CPC+, making it key for understanding impacts under a scale-up.

- **Compared with the evaluation sample.** The nationwide projected scaled sample was similar to the evaluation sample with respect to the only strong effect modifier (SSP status) identified in CART analysis, thus suggesting effects under a nationwide scale-up would likely not differ markedly from effects in the evaluation sample.

  The two samples differed on factors that did not moderate effects (and thus will not lead to differential impacts). Namely, compared with evaluation sample practices, nationwide projected scaled sample practices had less experience with primary care transformations (35 versus 47 percent), served slightly more disadvantaged populations (17 versus 15 percent of beneficiaries were dually eligible), and had different geographic distributions (e.g., 25 versus 15 percent in the South and 31 versus 38 percent in the Midwest).

For both Tracks 1 and 2, the nationwide projected scaled sample was similar to the evaluation sample with respect to the main effect modifier (SSP status), thus suggesting effects under a nationwide scale-up would likely not differ markedly from effects in the evaluation sample.

Similar differences hold in Track 2:

- **Compared with non-volunteering eligible practices.** We estimated the nationwide projected scaled sample practices would have more primary care transformation experience than non-volunteering practices (58 versus 18 percent), be more likely to be owned or managed by a health system or hospital (60 versus 30 percent), be more likely to have attested to meaningful use of an EHR earlier (80 versus 49 percent), and have more primary care practitioners (5.5 versus 3.2).
- **Compared with the evaluation sample.** Again, the national projected scaled sample did not differ meaningfully from the evaluation sample across the key effect modifier of SSP participation (46 versus 42 percent)[118]; we would therefore expect evaluation sample and projected scaled sample impacts to be similar.

  We estimated differences on factors that did not moderate effects: the nationwide projected scaled sample had less experience with primary care transformations (58 versus 73 percent), served slightly more disadvantaged populations (17 versus 14 percent of beneficiaries were dually eligible), and had different geographic distributions (e.g., 29 versus 15 percent in the South and 23 versus 38 percent in the Midwest).

[118] Although SSP was the strongest effect modifier in Track 2, it was a weaker effect modifier than in Track 1 and improved impact explainability by less than 10 percent.

Table 5.I.3. Comparison of characteristics of the evaluation sample, nationwide projected scaled sample, and eligible practices not in the nationwide projected scaled sample, with the key factor driving differences in impacts between practices in bold

Characteristic	Track 1				Track 2			
	A. Evaluation sample (n = 1,373)	B. Projected scaled sample (n ≈ 4,927)[a]	C. Eligible practices not in the projected scaled sample (n ≈ 43,819)[a]	Difference (B–C)	A. Evaluation sample (n = 1,515)	B. Projected scaled sample (n ≈ 5,160)[a]	C. Eligible practices not in the projected scaled sample (n ≈ 43,586)[a]	Difference (B–C)
Practice characteristics								
Number of primary care practitioners, mean[b]	4.2 (0.1)	4.5 (0.0)	3.2 (0.0)	1.3 (0.0)	5.2 (0.2)	5.5 (0.1)	3.2 (0.0)	2.3 (0.0)
Practice has nurse practitioners or physician assistants, %	50.4 (1.3)	52.5 (0.3)	41.7 (0.2)	10.7 (0.3)	58.8 (1.3)	60.5 (0.4)	41.7 (0.2)	18.7 (0.4)
Multispecialty practice, %	12.2 (0.9)	14.0 (0.2)	10.8 (0.1)	3.1 (0.2)	15.5 (0.9)	17.8 (0.3)	10.8 (0.1)	7.0 (0.2)
Owned (or managed) by a health system or hospital, %	52.7 (1.3)	53.4 (0.3)	29.9 (0.2)	23.5 (0.3)	56.8 (1.3)	59.7 (0.4)	29.9 (0.2)	29.7 (0.3)
Participation in a Medicare SSP ACO as of January 1, 2017, %	**53.8 (1.3)**	**54.6 (0.3)**	**33.1 (0.2)**	**21.5 (0.3)**	**42.0 (1.3)**	**46.4 (0.4)**	**33.1 (0.2)**	**13.3 (0.3)**
Selected primary care transformation experience, %[c]	47.3 (1.3)	34.7 (0.3)	18.4 (0.2)	16.4 (0.2)	73.3 (1.1)	58.0 (0.4)	18.4 (0.1)	39.6 (0.3)
EHR meaningful adoption 2011–2012, %[d]	71.5 (1.2)	71.3 (0.3)	48.6 (0.2)	22.7 (0.3)	82.2 (1.0)	79.7 (0.3)	48.6 (0.2)	31.1 (0.4)
Characteristics of Medicare FFS beneficiaries assigned to practices at baseline (2016)								
Number of assigned Medicare FFS beneficiaries at baseline, mean	638.9 (16.1)	666.6 (4.9)	416.9 (2.2)	249.7 (3.0)	706.3 (17.6)	716.9 (6.5)	416.9 (2.1)	300.0 (3.9)
Medicare expenditures per beneficiary ($/month), mean	887.2 (5.6)	897.7 (1.3)	978.5 (1.8)	-80.8 (2.1)	878.6 (5.2)	907.7 (1.7)	978.5 (1.8)	-70.8 (2.8)
Acute care stays per 1,000 beneficiaries (annualized), mean	288.5 (2.4)	283.0 (0.5)	309.3 (0.6)	-26.3 (0.7)	286.4 (2.1)	283.4 (0.6)	309.3 (0.6)	-26.0 (1.0)
ED visits per 1,000 beneficiaries (annualized), mean	512.3 (6.1)	512.3 (1.3)	556.7 (1.4)	-44.4 (1.6)	510.0 (5.3)	511.3 (1.5)	556.7 (1.4)	-45.4 (2.1)
Normalized HCC score among beneficiaries assigned in the baseline year, mean[e]	1.0 (0.0)	1.0 (0.0)	1.1 (0.0)	0.0 (0.0)	1.0 (0.0)	1.0 (0.0)	1.1 (0.0)	0.0 (0.0)
Beneficiary age, mean	71.3 (0.1)	71.5 (0.0)	70.9 (0.0)	0.6 (0.0)	71.0 (0.1)	71.2 (0.0)	70.9 (0.0)	0.4 (0.0)
Percentage of beneficiaries with Black race, %	6.5 (0.3)	7.2 (0.1)	10.8 (0.1)	-3.6 (0.1)	7.1 (0.3)	7.8 (0.1)	10.8 (0.1)	-3.0 (0.1)
Percentage of beneficiaries with White race, %	85.7 (0.5)	85.2 (0.1)	79.7 (0.1)	5.5 (0.1)	85.6 (0.5)	84.7 (0.1)	79.7 (0.1)	5.1 (0.2)
Percentage of beneficiaries with who are male, %	41.3 (0.2)	41.6 (0.1)	42.3 (0.0)	-0.7 (0.1)	41.9 (0.2)	41.8 (0.1)	42.3 (0.0)	-0.5 (0.1)
Percentage of beneficiaries with age as original reason for Medicare entitlement, %	77.8 (0.3)	78.5 (0.1)	75.2 (0.1)	3.4 (0.1)	77.9 (0.3)	78.1 (0.1)	75.2 (0.1)	2.9 (0.1)

Table 5.I.3. ***(continued)***

Characteristic	Track 1				Track 2			
	A. Evaluation sample (n = 1,373)	B. Projected scaled sample (n ≈ 4,927)[a]	C. Eligible practices not in the projected scaled sample (n ≈ 43,819)[a]	Difference (B–C)	A. Evaluation sample (n = 1,515)	B. Projected scaled sample (n ≈ 5,160)[a]	C. Eligible practices not in the projected scaled sample (n ≈ 43,586)[a]	Difference (B–C)
Percentage of beneficiaries who were dually eligible, %[f]	14.8 (0.4)	16.8 (0.1)	22.0 (0.1)	-5.2 (0.1)	14.4 (0.3)	16.9 (0.1)	22.0 (0.1)	-5.1 (0.2)
Characteristics of practices' geographic location								
South region, %[g]	14.9 (1.0)	25.3 (0.3)	39.9 (0.2)	-14.6 (0.3)	15.2 (0.9)	28.8 (0.3)	39.9 (0.2)	-11.1 (0.4)
Midwest region, %[g]	37.7 (1.3)	31.0 (0.3)	20.5 (0.2)	10.4 (0.2)	37.7 (1.2)	23.3 (0.3)	20.5 (0.2)	2.8 (0.3)
Northeast region, %[g]	30.8 (1.2)	24.6 (0.3)	21.1 (0.2)	3.5 (0.2)	28.0 (1.2)	26.7 (0.4)	21.1 (0.2)	5.7 (0.3)
West region, %[g]	16.6 (1.0)	19.2 (0.3)	18.5 (0.2)	0.7 (0.2)	19.1 (1.0)	21.2 (0.3)	18.5 (0.2)	2.6 (0.3)
HRR price index (measure of relative costs in the HRR)	1.1 (0.0)	1.1 (0.0)	1.1 (0.0)	0.0 (0.0)	1.1 (0.0)	1.1 (0.0)	1.1 (0.0)	0.0 (0.0)
Median household income in the county where the practice is located	58,118.1 (420.5)	58,077.5 (99.8)	56,048.6 (70.7)	2,028.9 (87.4)	57,656.5 (373.8)	57,972.6 (106.5)	56,048.6 (71.8)	1,924.0 (113.2)
Suburban location, %	17.4 (1.0)	18.9 (0.3)	14.4 (0.2)	4.4 (0.2)	12.3 (0.8)	14.8 (0.3)	14.4 (0.2)	0.3 (0.3)
Urban location, %	72.8 (1.2)	71.8 (0.3)	75.5 (0.2)	-3.7 (0.2)	79.7 (1.0)	78.0 (0.3)	75.5 (0.2)	2.5 (0.3)
Rural location, %	9.8 (0.8)	9.4 (0.2)	10.1 (0.1)	-0.7 (0.2)	7.9 (0.7)	7.2 (0.2)	10.1 (0.1)	-2.9 (0.2)
Percentage of people 25 or older in practice county with 4 years of college education	30.9 (0.3)	30.5 (0.1)	29.4 (0.0)	1.1 (0.1)	31.1 (0.3)	30.9 (0.1)	29.4 (0.0)	1.5 (0.1)
Percentage of residents in practice county below poverty level in 2014	13.9 (0.1)	14.2 (0.0)	15.5 (0.0)	-1.3 (0.0)	14.1 (0.1)	14.3 (0.0)	15.5 (0.0)	-1.2 (0.0)
Number of beds per population	30.4 (0.5)	30.7 (0.1)	31.2 (0.1)	-0.5 (0.1)	30.3 (0.4)	29.8 (0.2)	31.2 (0.1)	-1.4 (0.2)
Medicare Advantage penetration rate, %	28.9 (0.3)	27.7 (0.1)	30.7 (0.1)	-3.1 (0.1)	32.8 (0.3)	32.1 (0.1)	30.7 (0.1)	1.4 (0.1)
COVID-19 characteristics								
2021 Wave 3 pandemic vulnerability index[h]	0.51 (0.05)	0.51 (0.05)	0.53 (0.05)	-0.01 (<0.01)	0.51 (0.05)	0.52 (0.05)	0.53 (0.05)	-0.01 (<0.01)
2021 Wave 5 pandemic vulnerability index[h]	0.50 (0.05)	0.50 (0.05)	0.51 (0.05)	-0.01 (<0.01)	0.50 (0.05)	0.50 (0.05)	0.51 (0.05)	-0.01 (<0.01)
2021 Wave 3 excess deaths[h]	6.54 (3.31)	7.88 (4.56)	9.30 (4.95)	-1.42 (0.04)	6.92 (3.47)	8.25 (4.79)	9.26 (4.92)	-1.00 (0.06)
Peak 2021 excess deaths[h]	12.91 (4.01)	14.09 (5.89)	15.97 (6.54)	-1.88 (0.05)	13.50 (3.90)	14.66 (6.08)	15.89 (6.54)	-1.23 (0.08)
Social vulnerability index[h]	0.43 (0.26)	0.46 (0.27)	0.51 (0.27)	-0.05 (<0.01)	0.42 (0.26)	0.46 (0.27)	0.51 (0.27)	-0.06 (<0.01)
Government response index[h]	53.53 (3.53)	52.83 (4.72)	52.33 (5.02)	0.50 (0.04)	53.82 (3.90)	53.05 (4.84)	52.27 (5.00)	0.78 (0.06)

The table presents proportion or mean (SE). Sample sizes correspond to numbers of practices.

[a] Approximate sample sizes based on the propensity for volunteering. Nationwide projected scaled sample characteristics are weighted averages of national practice characteristics, weighted by the propensity for volunteering. Characteristics for eligible practices not in the nationwide projected scaled sample are weighted by one minus the propensity for volunteering.

Table 5.I.3. ***(continued)***

[b] We defined primary care practitioners using practitioner specialty information from NPPES and SK&A data. For practitioners with a valid NPI, we identified a practitioner as primary care using primary and secondary taxonomy codes in the NPPES (following the approach used in CPC+ payment methodology); for practitioners without an NPI in the SK&A data, we identified a practitioner as primary care using practitioner specialty information from SK&A (practitioner specialty was either family practice, general practice, geriatrician, internist, or internist and pediatrics).

[c] Percentage of practices that hold NCQA, TJC, AAAHC, URAC, or state medical-home recognition, or have participated in CPC Classic, CMMI's Transforming Clinical Practice Initiative, or CMMI's Multi-Payer Advanced Primary Care Program as of 2014.

[d] At least one practitioner attested to meaningful use under the Medicare EHR Incentive Program from 2011 to 2015.

[c] The (baseline) 2016 HCC score is based on beneficiaries' diagnoses in 2015.

[f] Calculated as the percentage of beneficiaries assigned to a practice in the baseline year who were dually eligible for Medicare and Medicaid in the quarter before the start of the baseline year.

[g] U.S. census region.

[h] See Appendix 5.D for definitions of pandemic waves in 2021, excess deaths, the pandemic vulnerability index, and the social vulnerability index.

AAAHC = Accreditation Association for Ambulatory Health Care; ACO = Accountable Care Organization; CMMI = Center for Medicare & Medicaid Innovation; CPC+ = Comprehensive Primary Care Plus; ED = emergency department; EHR = electronic health record; FFS = fee-for-service; HCC = hierarchical condition category; HRR = hospital referral region; NCQA = National Committee for Quality Assurance; NPI = National Provider Identifier; NPPES = National Plan and Provider Enumeration System; SE = standard error; SSP = Medicare Shared Savings Program; TJC = the Joint Commission; URAC = Utilization Review Accreditation Commission.

Steps 2 and 3. CPC+ impact estimates in the evaluation sample and under two scale-up scenarios for Tracks 1 and 2

In this section, we present findings using PY 5 estimates for (1) Medicare expenditures excluding enhanced CPC+ payments, (2) outpatient ED visits including observation stays, and (3) acute hospitalizations. For each of these three outcomes and for each track, we present impact estimates for the overall (nationwide) projected scaled sample and targeted subsets of practices most likely to generate savings.

A. Medicare expenditures excluding enhanced CPC+ payments

Key takeaways

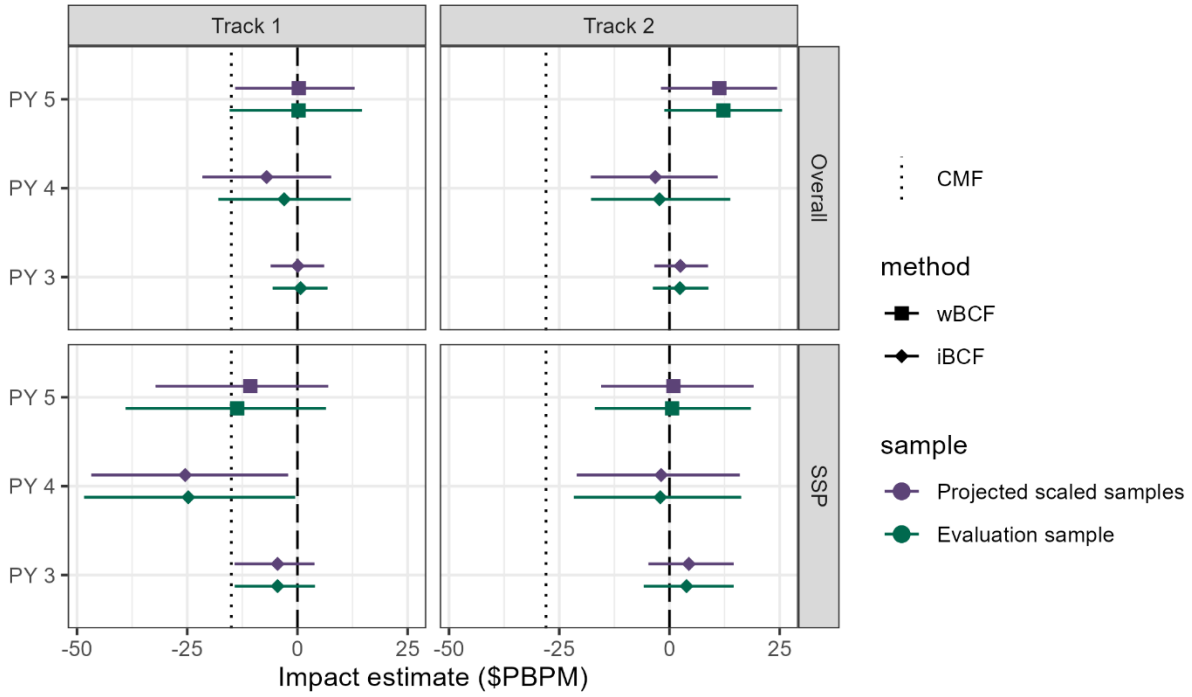
No Track 1 or 2 CPC+ scale-up (nationwide or targeted to practices most likely to benefit) is likely to generate sufficient savings to offset CMFs. A scale-up of Track 1 targeted to baseline SSP practices is estimated to have the most favorable impacts: 34 percent probability of cost neutrality, 82 percent probability of decreasing outpatient ED visits, and 89 percent probability of decreasing acute hospitalizations.

Scaling up the model nationally is unlikely to result in savings. As expected based on the similarity of effect modifiers between the nationwide projected scaled sample and evaluation sample (Table 5.I.3), wBCF-based estimates of CPC+ effects on Medicare expenditures without enhanced payments were similar for the nationwide projected scaled sample and the evaluation sample (Table 5.I.4). A nationwide scale-up was estimated to have null impacts for Track 1 (90 percent credible interval [CI] -$14 to $13 PBPM) and to have a 92 percent probability of increasing expenditures for Track 2 (mean projected increase excluding CMFs of $11 PBPM, 90 percent CI -$2 to $24; Table 5.I.4, Figure 5.I.2).

No Track 1 or 2 targeting approach is likely to generate savings. However, a scale-up of Track 1 to SSP practices would likely show the most favorable impacts. In PY 5, as in PY 4 and PY 3, participation in SSP (based on the definition of SSP as of January 1, 2017) was the key driver of favorable effects for Track 1. SSP practices represent 55 percent of practices volunteering for a scale-up of Track 1 (Table 5.I.3). However, the estimated magnitude of savings among SSP practices attenuated in PY 5 compared with PY 4. As a result, there was an estimated probability of only 34 percent that a scale-up targeted to SSP practices would be cost neutral (-$11 PBPM, 90 percent CI -$32 to $7), down from an estimated 79 percent probability in PY 4 (-$25 PBPM, 90 percent CI -$47 to -$2). The Discussion section considers possible reasons for this attenuation.

A nationwide scale-up of Track 2 was estimated to lead to null effects on expenditures for SSP practices (90 percent CI -$16 to $19 PBPM). Variation in CPC+ effects was not strongly driven by any observed covariates for Track 2; namely, no factor improved CART's ability to explain impacts by at least 10 percent. However, among the less-explanatory factors, SSP participation was the strongest effect modifier according to CART.

Figure 5.I.2. Estimated impacts of CPC+ on Medicare expenditures, excluding enhanced CPC+ payments in the projected scaled samples in PYs 3 to 5 (estimate and 90 percent CI)



Note: Each point represents the estimated impact on Medicare expenditures and each horizontal line represents the corresponding 90 percent credible interval. The dashed vertical line marks a $0 PBPM impact, and the dotted vertical line corresponds to an impact equal in size to the average projected CMF nationwide ($15 PBPM for Track 1 and $28 PBPM for Track 2).

CI = credible interval; CMF = care management fee; iBCF = individualized weighted Bayesian Causal Forest; PBPM = per beneficiary per month; PY = Program Year; SSP = Medicare Shared Savings Program; wBCF = weighted Bayesian Causal Forest.

BCF estimates of uncertainty increased during COVID-19 years (PY 4 and PY 5), relative to PY 3. Compared with PY 3 impact estimates, BCF estimated greater uncertainty around all PY 4 and PY 5 impact estimates; the DD estimates did not experience such an increase in uncertainty (Figure 5.I.2; Table 5.I.4). See Sensitivity Analysis C for further discussion of this finding.

Table 5.I.4. Estimated CPC+ impacts and probabilities of reducing Medicare expenditures in the evaluation sample and projected scaled samples in PY 3 to PY 5

		Track 1				Track 2			
	Subgroup	Estimated impact on Medicare expenditures excluding enhanced CPC+ payments (90% CI) $PBPM	Estimated probability of reducing Medicare expenditures excluding enhanced CPC+ payments (%)	Probability of sufficient reduction to offset CMFs (%)	Aggregate annual impact estimates for Medicare expenditures including CMFs (90% CI) million $[a]	Estimated impact on Medicare expenditures excluding enhanced CPC+ payments (90% CI) $PBPM	Estimated probability of reducing Medicare expenditures excluding enhanced CPC+ payments (%)	Probability of sufficient reduction to offset CMFs (%)	Aggregate annual impact estimates for Medicare expenditures including CMFs (90% CI) million $[a]
PY 5 (wBCF)									
Projected scaled samples	Overall	0 (-14, 13)	47	4	601 (36, 1,103)	11 (-2, 24)	8	<1	1,754 (1,150, 2,383)
	SSP	-11 (-32, 7)	80	34	93 (-373, 485)	1 (-16, 19)	47	1	620 (268, 1,015)
Evaluation sample	Overall	0 (-15, 15)	47	9	132 (-33, 283)	12 (-1, 26)	6	<1	461 (289, 631)
	SSP	-14 (-30, 6)	84	50	-8 (-145, 101)	1 (-17, 18)	48	1	137 (38, 239)
PY 4 (iBCF)									
Projected scaled samples	Overall	-7 (-22, 8)	78	18	315 (-251, 887)	-3 (-18, 11)	64	<1	1,106 (447, 1,736)
	SSP	-25 (-47, -2)	96	79	-228 (-687, 291)	-2 (-21, 16)	56	1	560 (148, 949)
Evaluation sample	Overall	-3 (-18, 12)	63	15	97 (-60, 257)	-2 (-18, 14)	60	1	274 (76, 480)
	SSP	-25 (-48, -1)	95	80	-68 (-196, 64)	-2 (-22, 16)	57	4	122 (11, 226)
PY 3 (iBCF)									
Projected scaled samples	Overall	0 (-6, 6)	47	<1	592 (349, 831)	2 (-3, 9)	25	<1	1,359 (1,074, 1,659)
	SSP	-5 (-14, 4)	78	4	229 (16, 416)	4 (-5, 15)	23	<1	695 (500, 919)
Evaluation sample	Overall	1 (-6, 7)	42	<1	141 (75, 205)	2 (-4, 9)	26	<1	342 (264, 425)
	SSP	-5 (-14, 4)	79	8	44 (-8, 90)	4 (-6, 15)	27	<1	159 (105, 220)

[a] Calculated as the estimated PBPM impact on Medicare expenditures including CMFs multiplied by the estimated number of Medicare FFS beneficiaries assigned to a practice in the baseline year and by 12 months. Note that, because beneficiary assignment and eligibility were not available for the projected scaled samples, to calculate aggregate annual impact estimates, we instead assumed all beneficiaries assigned at baseline would be eligible for the full 12 months of each PY. For consistency with the projected scaled sample methodology, we likewise used this approach for the evaluation sample rather than using eligible beneficiary months.

CI = credible interval; CMF = care management fee; CPC+ = Comprehensive Primary Care Plus; DD = difference-in-differences; FFS = fee-for-service; iBCF = individualized weighted Bayesian Causal Forest; PBPM = per beneficiary per month; PY = Program Year; SSP = Medicare Shared Savings Program; wBCF = weighted Bayesian Causal Forest.

Table 5.I.5. Sample sizes for evaluation sample and projected scaled samples

Subgroup	Track 1				Track 2			
	Estimated number of participating practices[a]	Estimated number of assigned Medicare FFS beneficiaries[b]	Proportion of eligible practices nationwide (%)	Proportion of assigned Medicare FFS beneficiaries in eligible practices nationwide (%)	Estimated number of participating practices[a]	Estimated number of assigned Medicare FFS beneficiaries[b]	Proportion of eligible practices nationwide (%)	Proportion of assigned Medicare FFS beneficiaries in eligible practices nationwide (%)
Projected scaled samples								
Overall	4,927	3,270,535	10	16	5,160	3,715,291	11	18
SSP	2,670	1,816,246	5	9	2,390	1,787,113	5	9
Evaluation sample								
Overall	1,373	877,150	3	4	1,515	1,070,048	3	5
SSP	738	451,485	2	2	636	472,404	1	2
Non-SSP	635	425,665	1	2	879	597,645	2	3

[a] For the evaluation sample: the number of CPC+ practices that began their CPC+ participation in 2017 for the subgroup. For the projected scaled samples: the sum across practices' propensities for volunteering, for the subgroup.

[b] For the evaluation sample: the number of beneficiaries assigned to a practice in the baseline year, summed over the subgroup. For the projected scaled sample: the number of beneficiaries assigned to a practice in the baseline year multiplied by the propensity for volunteering for that practice, summed over the subgroup.

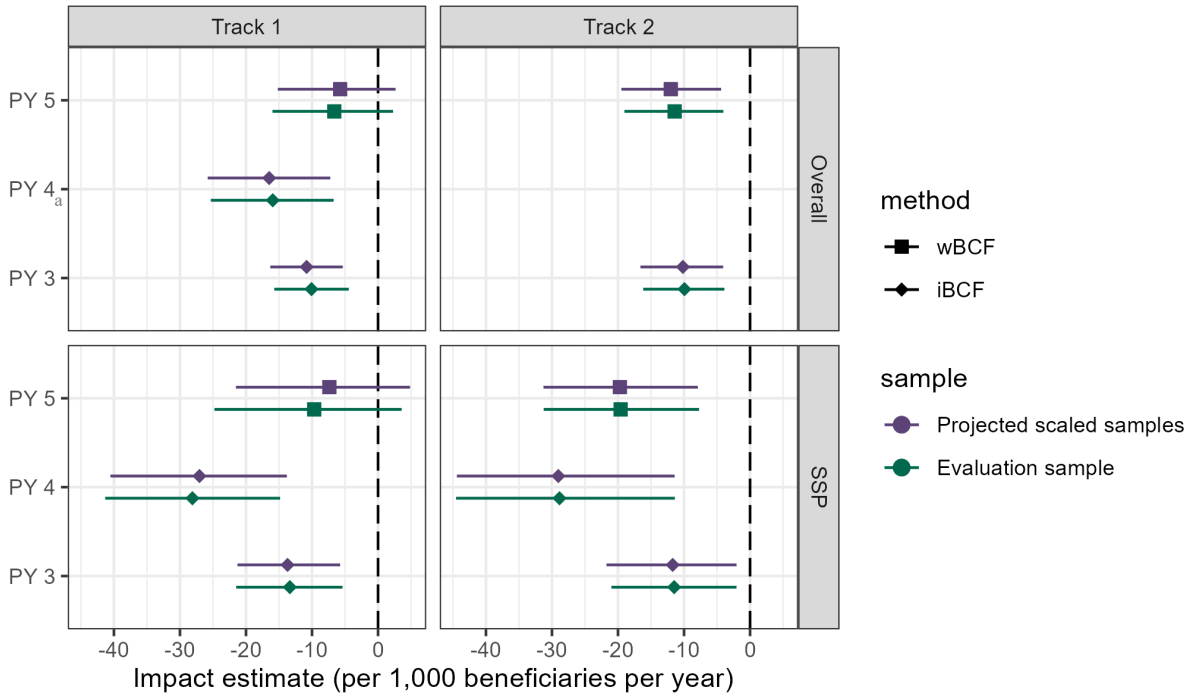
B. Service use

Outpatient ED visits, including observation stays

Across Track 1 and Track 2, both a nationwide scale-up and a scale-up targeted to SSP practices is likely to decrease outpatient ED visits. Expanding Track 1 would have an 86 percent probability of decreasing outpatient ED visits for a nationwide scale-up (Figure 5.I.3; Table 5.I.6) and an 82 percent probability of reductions for a targeted scale-up to SSP practices. Similarly, scaling up Track 2 would have > 99 percent probability of decreasing outpatient ED visits for both nationwide and SSP-targeted scale-ups.

Favorable PY 4 outpatient ED visit impact estimates for Track 1 and 2 SSP practices attenuated in PY 5. For example, for Track 1 SSP practices, we estimated a reduction of 7 per 1,000 beneficiaries per year (90 percent CI -22 to 5) based on PY 5 outcomes compared with a reduction of 28 per 1,000 beneficiaries per year (90 percent CI -41 to -15) based on PY 4 outcomes.

Figure 5.I.3. Estimated impacts of CPC+ on outpatient ED visits including observation stays in the projected scaled samples in PYs 3 to 5 (estimate and 90% CI)



Note: Each dot represents the estimated impact on annualized outpatient ED visits including observation stays, per 1,000 beneficiaries. Each horizontal line represents the corresponding 90 percent credible interval. The dashed vertical line marks an impact estimate of zero.

[a] Results are not available for Track 2 overall PY 4 estimates as the iBCF model did not converge for non-SSP practices, whose estimates are used in overall estimates.

CI = credible interval; ED = emergency department; iBCF = individualized weighted Bayesian Causal Forest; PY = Program Year; SSP = Medicare Shared Savings Program; wBCF = weighted Bayesian Causal Forest.

Table 5.I.6. Estimated impacts of CPC+ and probabilities of reducing outpatient ED visits including observation stays in the projected scaled samples in PY 3 to PY 5

	Subgroup	Estimated impact per 1,000 beneficiaries per year (90% CI)	Estimated probability of reducing ED visits (%)	Aggregate annual impact estimates (thousands)[a]	Estimated impact per 1,000 beneficiaries per year (90% CI)	Estimated probability of reducing ED visits (%)	Aggregate annual impact estimates (thousands)[a]
PY 5 (wBCF)							
Projected scaled samples	Overall	-6 (-15, 3)	86	-19 (-50, 9)	-12 (-19, -4)	>99	-45 (-72, -16)
	SSP	-7 (-22, 5)	82	-13 (-39, 9)	-20 (-31, -8)	>99	-35 (-56, -14)
Evaluation sample	Overall	-7 (-16, 2)	89	-6 (-14, 2)	-11 (-19, -4)	99	-12 (-20, -4)
	SSP	-10 (-25, 4)	87	-4 (-11, 2)	-20 (-31, -8)	>99	-9 (-15, -4)
PY 4 (iBCF)							
Projected scaled samples	Overall	-18 (-28, -8)	>99	-58 (-91, -26)	-13 (-22, -3)	97	-47 (-84, -9)
	SSP	-28 (-41, -15)	>99	-51 (-76, -27)	-27 (-42, -9)	>99	-47 (-74, -17)
Evaluation sample	Overall	-17 (-27, -8)	>99	-15 (-23, -7)	NA[b]	NA[b]	NA[b]
	SSP	-29 (-42, -16)	>99	-13 (-19, -7)	-27 (-42, -10)	99	-13 (-20, -5)
PY 3 (iBCF)							
Projected scaled samples	Overall	-11 (-16, -5)	>99	-35 (-54, 18)	-10 (-17, -4)	>99	-38 (-62, -15)
	SSP	-14 (-21, -6)	>99	-25 (-39, -10)	-12 (-22, -2)	98	-21 (-40, -4)
Evaluation sample	Overall	-10 (-16, -4)	>99	-9 (-14, -4)	-10 (-16, -4)	>99	-11 (-17, -4)
	SSP	-13 (-21, -5)	99	-6 (-10, -2)	-11 (-21, -2)	98	-5 (-10, -1)

[a] Calculated as the estimated impact per year multiplied by the estimated number of assigned Medicare FFS beneficiaries. Note that, because beneficiary attribution and eligibility were not available for the projected scaled samples, to calculate aggregate annual impact estimates, we instead assumed all beneficiaries assigned at baseline would be eligible for the full 12 months of each program year. For consistency with the projected scaled sample methodology, we likewise used this approach for the evaluation sample rather than using eligible beneficiary months.

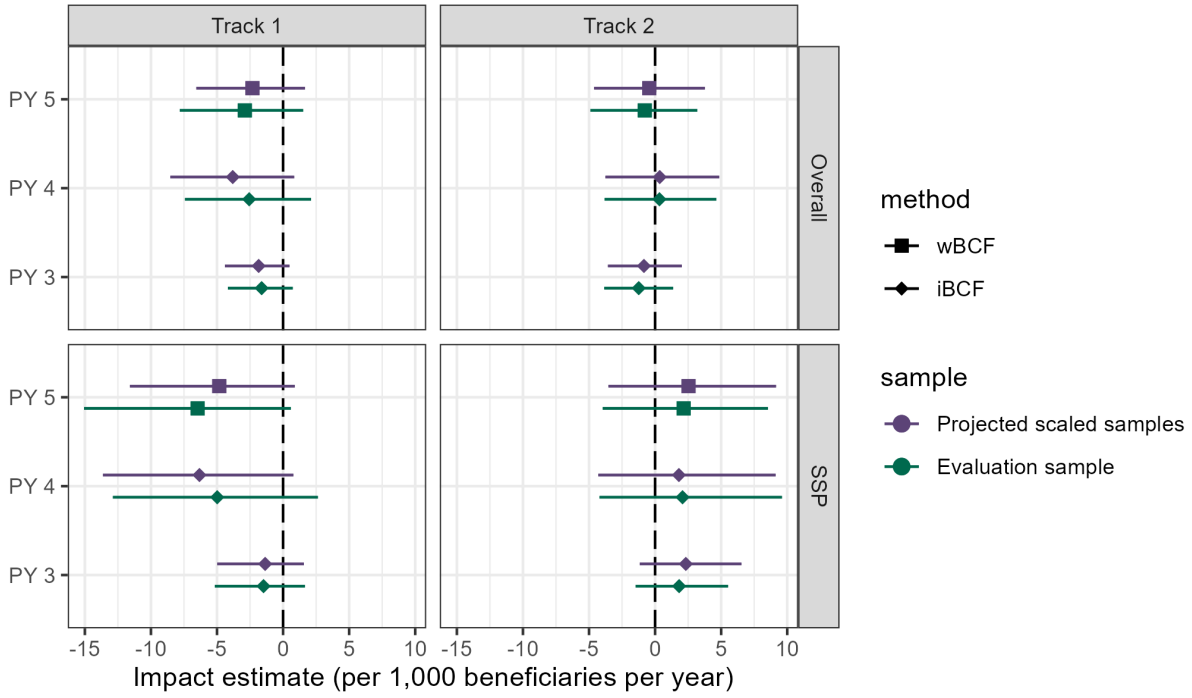
[b] Results are not available for Track 2 overall PY 4 estimates as the iBCF model did not converge for non-SSP practices, whose estimates are used in overall estimates.

CI = credible interval; DD = difference-in-differences; ED = emergency department; FFS = fee-for-service; iBCF = individualized weighted Bayesian Causal Forest; NA = not available; PY = Program Year; SSP = Medicare Shared Savings Program; wBCF = weighted Bayesian Causal Forest.

Acute hospitalizations

Scaling up Track 1 nationwide has an 81 percent probability of reducing acute hospitalizations, with higher probability of reductions for an SSP-targeted scale-up. Namely, a nationwide Track 1 scale-up is estimated to reduce annual hospitalizations by 2 per 1,000 beneficiaries (90 percent CI -7 to 2; Figure 5.I.4Figure 5.I.7, Table 5.I.7). Meanwhile, a Track 1 scale-up targeting SSP practices is estimated to have an 89 percent probability of reducing hospitalizations for an estimated annual reduction of 5 per 1,000 beneficiaries (90 percent CI -12 to 1). **However, scaling up Track 2 is less likely to reduce acute hospitalizations than scaling up Track 1, regardless of whether the scale-up is nationwide (57 versus 81 percent probability) or targeted to practices participating in SSP (26 versus 89 percent probability).** Acute hospitalization impact estimates remained consistent between PY 4 and PY 5.

Figure 5.I.4. Estimated impacts of CPC+ on acute hospitalizations in the projected scaled samples in PY 3 to PY 5 (estimate and 90% CI)



Note: Each dot represents the estimated impact on annualized acute hospitalizations, per 1,000 beneficiaries. Each horizontal line represents the corresponding 90 percent credible interval. The dashed vertical line marks an impact estimate of zero.

CI = credible interval; iBCF = individualized weighted Bayesian Causal Forest; PY = Program Year; SSP = Medicare Shared Savings Program; wBCF = weighted Bayesian Causal Forest.

Table 5.I.7. Estimated impacts of CPC+ and probabilities of reducing acute hospitalizations in the evaluation sample and projected scaled samples in PY 3 to PY 5

		Track 1			Track 2		
	Subgroup	Estimated impact per 1,000 beneficiaries per year (90% CI)	Estimated probability of reducing acute hospitalizations (%)	Aggregate annual estimated impact (thousands)[a]	Estimated impact per 1,000 beneficiaries per year (90% CI)	Estimated probability of reducing acute hospitalizations (%)	Aggregate annual estimated impact (thousands)[a]
PY 5 (wBCF)							
Projected scaled samples	Overall	-2 (-7, 2)	81	-8 (-22, 6)	0 (-5, 4)	57	-2 (-17, 14)
	SSP	-5 (-12, 1)	89	-9 (-21, 2)	3 (-4, 9)	26	5 (-6, 16)
Evaluation sample	Overall	-3 (-8, 2)	84	-3 (-7, 1)	-1 (-5, 3)	62	-1 (-5, 3)
	SSP	-6 (-15, 1)	92	-3 (-7, 0)	2 (-4, 9)	29	1 (-2, 4)
PY 4 (iBCF)							
Projected scaled samples	Overall	-4 (-9, 1)	91	-12 (-28, 3)	0 (-4, 5)	45	1 (-14, 18)
	SSP	-6 (-14, 1)	93	-11 (-25, 1)	2 (-4, 9)	33	3 (-8, 16)
Evaluation sample	Overall	-3 (-7, 2)	81	-2 (-7, 2)	0 (-4, 5)	45	0 (-4, 5)
	SSP	-5 (-13, 3)	86	-2 (-6, 1)	2 (-4, 10)	32	1 (-2, 5)
PY 3 (iBCF)							
Projected scaled samples	Overall	-2 (-4, 0)	91	-6 (-15, 2)	-1 (-4, 2)	70	-3 (-13, 8)
	SSP	-1 (-5, 2)	74	-2 (-9, 3)	2 (-1, 7)	16	4 (-2, 12)
Evaluation sample	Overall	-2 (-4, 1)	87	-1 (-4, 1)	-1 (-4, 1)	78	-1 (-4, 1)
	SSP	-1 (-5, 2)	76	-1 (-2, 1)	2 (-1, 6)	20	1 (-1, 3)

[a] Calculated as the estimated impact per year multiplied by the estimated number of assigned Medicare FFS beneficiaries. Note that, because beneficiary attribution and eligibility were not available for the projected scaled samples, to calculate aggregate annual impact estimates, we instead assumed all beneficiaries assigned at baseline would be eligible for the full 12 months of each program year. For consistency with the projected scaled sample methodology, we likewise used this approach for the evaluation sample rather than using eligible beneficiary months.

CI = credible interval; FFS = fee-for-service; iBCF = individualized weighted Bayesian Causal Forest; PY = Program Year; SSP = Medicare Shared Savings Program; wBCF = weighted Bayesian Causal Forest.

Step 4. Assess sensitivity to assumptions

A. Sensitivity Analysis A: Unmeasured modifiers of effects on Medicare expenditures

Our conclusion—that a cost-neutral scale-up is unlikely—would not change in the face of bias from an unmeasured subgroup variable, such as practices' motivation to improve, as strong as the strongest or second-strongest measured effect modifier bias. For expenditures excluding enhanced CPC+ payments, such an unmeasured effect modifier would, in the more extreme case, *change projected scaled sample impacts by about $2 PBPM for Track 1 and $3 PBPM for Track 2 ($1 each in the less extreme case).* With such a shift, a scale-up of Track 1 to SSP practices could have a probability of offsetting CMFs as low as 28 percent or as high as 41 percent (compared with 34 percent in the main analysis).

B. Sensitivity Analysis B: Analytic approach comparison for each of the three key outcomes

To assess the sensitivity of findings to different confounding adjustment approaches, this analysis compared evaluation sample impact estimates from wBCF with those from the frequentist DD presented in Chapter 5. Across expenditure, outpatient ED visit, and acute hospitalization outcomes, evaluation sample impact estimates based on wBCF aligned with those from the main DD analysis presented in Chapter 5 for Track 1 (overall, SSP and non-SSP groups) and for Track 2 non-SSP practices (Table 5.I.8; Figure 5.I.5, Figure 5.I.6, Figure 5.I.7; note Figure 5.I.5 also includes results for Sensitivity Analysis C).

For example, for SSP practices that participated in Track 1, both analytic approaches estimated a decrease in expenditures, though wBCF estimated larger uncertainty bounds that subsumed DD's estimate and confidence interval: wBCF estimated a reduction of $14 PBPM (90 percent CI -$39 to $6) while DD estimated a reduction of $19 (90 percent CI -$30 to -$9). For non-SSP practices in Track 1, both approaches estimated increases in expenditures, with 90 percent CIs crossing zero. Similarly, both wBCF and DD estimate reductions in Track 1 ED visits for all groups. However, wBCF's uncertainty bounds exceed those from DD and crossed zero. Likewise, both wBCF and DD estimated reductions in acute hospitalizations for Track 1 SSP practices and null impacts for non-SSP practices. For Track 2 non-SSP practices, both approaches estimated increases in expenditures (though DD's CI crossed zero), no impacts on outpatient ED visits (though wBCF's estimates trended towards reductions), and null impacts on acute hospitalizations trending toward reductions.

For Track 2 SSP practices, expenditures estimates from the two approaches were less well aligned: wBCF estimated null expenditure effects (90 percent CI -$17 to $18 PBPM) while DD estimated savings (-$17 PBPM, 90 percent CI -$32 to -$2 PBPM).

These findings increase our confidence that CPC+ Track 1 impact estimates and Track 2 non-SSP *impact estimates are not overly sensitive to different approaches for confounding adjustment.* Track 2 SSP impact estimates show more sensitivity, as Sensitivity Analysis C shows, reinforcing our lack of confidence in a feasible Track 2 scale-up.

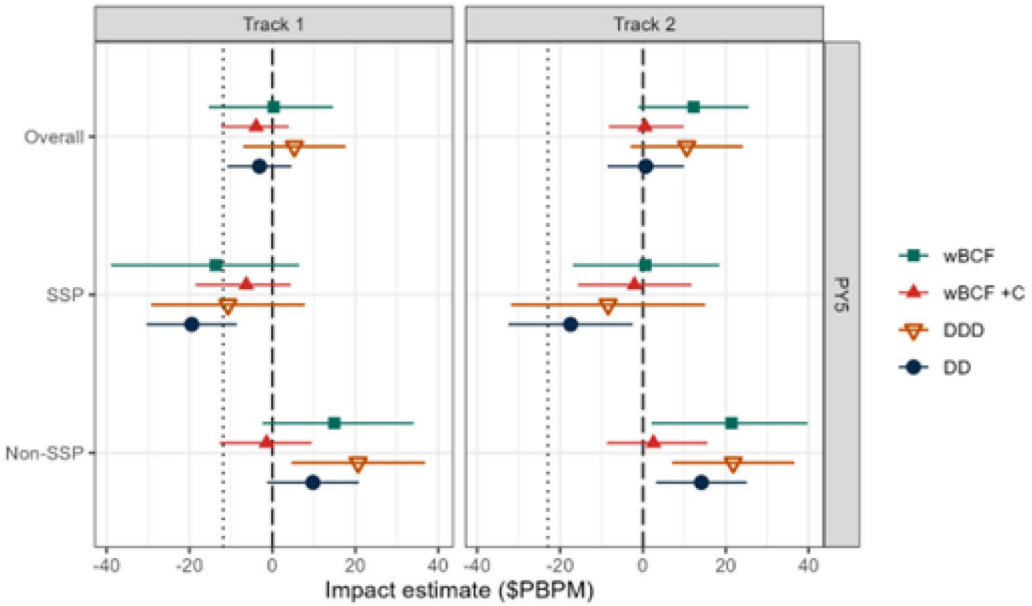
Table 5.I.8. Estimated impacts of CPC+ on Medicare expenditures excluding enhanced CPC+ payments, outpatient ED visits, and acute hospitalizations in the evaluation sample in PY 3 to PY 5 (estimate and 90% CI) when using wBCF and DD

		Estimated impact on Medicare expenditures excluding enhanced CPC+ payments (90% CI) $PBPM		Estimated outpatient ED visit impact per 1,000 beneficiaries per year (90% CI)		Estimated acute hospitalization impact per 1,000 beneficiaries per year (90% CI)	
	Subgroup	Track 1	Track 2	Track 1	Track 2	Track 1	Track 2
PY 5							
Method: wBCF	Overall	0 (-15, 15)	12 (-1, 26)	-7 (-16, 2)	-11 (-19, -4)	-3 (-8, 2)	-1 (-5, 3)
	SSP	-14 (-39, 6)	1 (-17, 18)	-10 (-25, 4)	-20 (-31, -8)	-6 (-15, 1)	2 (-4, 9)
	Non-SSP	15 (-2, 34)	21 (2, 40)	-3 (-15, 8)	-5 (-15, 4)	1 (-3, 5)	-3 (-9, 2)
Method: DD	Overall	-3 (-11, 5)	1 (-9, 10)	-16 (-22, -10)	-11 (-17, -5)	-3 (-6, 0)	-2 (-5, 1)
	SSP	-19 (-30, -9)	-17 (-32, -2)	-18 (-25, -10)	-20 (-29, -11)	-5 (-9, -1)	-2 (-7, 3)
	Non-SSP	10 (-1, 21)	14 (3, 25)	-14 (-23, -5)	-1 (-9, 6)	1 (-4, 5)	-2 (-6, 2)
PY 4							
Method: iBCF	Overall	-3 (-18, 12)	-2 (-18, 14)	-17 (-27, -8)	NA[a]	-3 (-7, 2)	0 (-4, 5)
	SSP	-25 (-48, -1)	-2 (-22, 16)	-29 (-42, -16)	-27 (-42, -10)	-5 (-13, 3)	2 (-4, 10)
	Non-SSP	20 (3, 39)	-2 (-26, 22)	-5 (-18, 9)	NA[a]	0 (-6, 6)	-1 (-6, 4)
Method: DD	Overall	-3 (-10, 5)	-2 (-11, 6)	-11 (-16, -5)	-8 (-14, -3)	-5 (-8, -2)	-5 (-8, -2)
	SSP	-15 (-26, -4)	-14 (-28, -1)	-14 (-21, -6)	-19 (-28, -11)	-8 (-12, -4)	-4 (-9, 1)
	Non-SSP	10 (0, 20)	9 (-1, 19)	-7 (-15, 2)	2 (-5, 10)	-1 (-6, 3)	-5 (-9, -1)
PY 3							
Method: iBCF	Overall	1 (-6, 7)	2 (-4, 9)	-10 (-16, -4)	-10 (-16, -4)	-2 (-4, 1)	-1 (-4, 1)
	SSP	-5 (-14, 4)	4 (-6, 15)	-13 (-21, -5)	-11 (-21, -2)	-1 (-5, 2)	2 (-1, 6)
	Non-SSP	6 (-2, 15)	1 (-6, 9)	-7 (-15, 1)	-9 (-17, -1)	-2 (-5, 2)	-4 (-7, 0)
Method: DD	Overall	2 (-5, 9)	-2 (-9, 6)	-8 (-13, -3)	-7 (-12, -3)	-3 (-6, 0)	-5 (-8, -2)
	SSP	-8 (-17, 1)	-8 (-20, 4)	-7 (-13, -1)	-8 (-15, -1)	-5 (-9, -1)	-2 (-7, 2)
	Non-SSP	14 (4, 24)	3 (-6, 13)	-9 (-16, -1)	-7 (-14, 0)	0 (-5, 5)	-7 (-11, -3)

[a] Results are not available for Track 2 overall PY 4 estimates as the iBCF model did not converge for non-SSP practices, whose estimates are used in overall estimates.

CI = confidence interval or credible interval; DD = difference-in-differences; ED = emergency department; iBCF = individualized weighted Bayesian Causal Forest; NA = not available; PY = Program Year; SSP = Medicare Shared Savings Program.

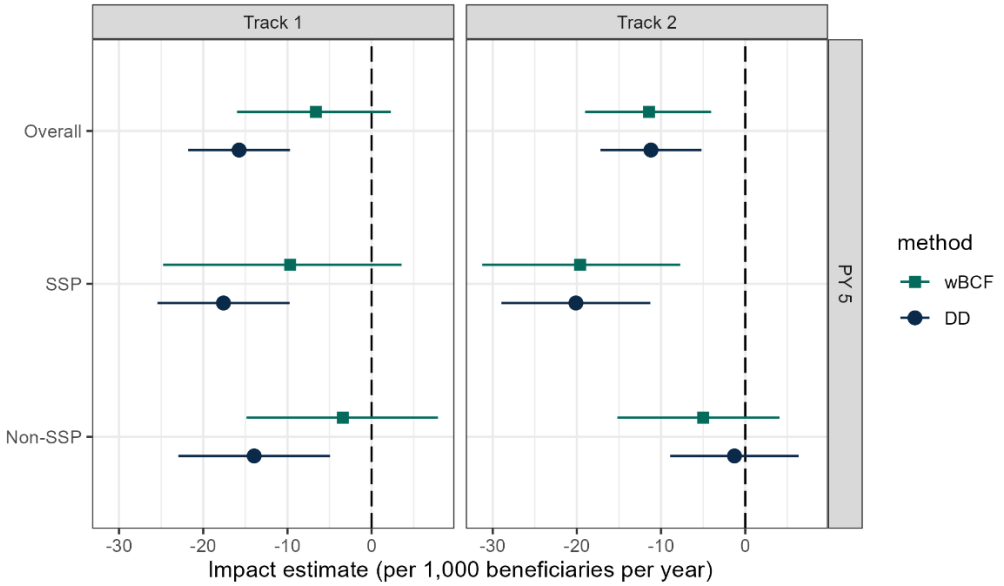
Figure 5.I.5. Estimated impacts of CPC+ on Medicare expenditures excluding enhanced CPC+ payments in the evaluation sample in PY 5 (estimate and 90 percent CI) when using wBCF, DD, wBCF expanded comparison (wBCF +C), and the frequentist DDD models



Note: Although the wBCF +C models did not fully converge, results aligned with those from a BART +C models, which did converge.

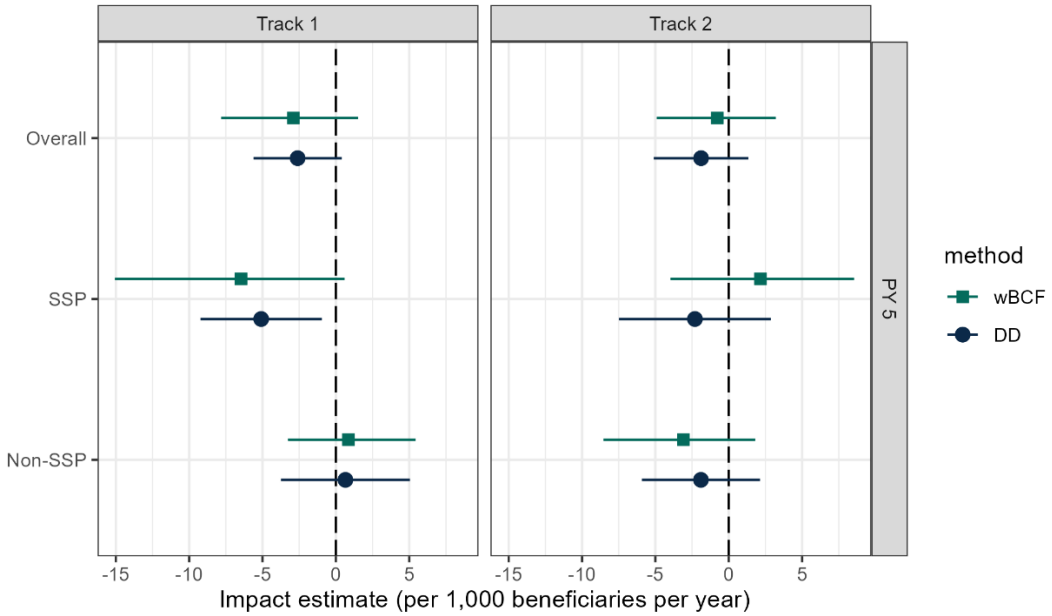
CI = confidence interval or credible interval; DD = difference-in-differences; DDD = triple differences; PY = Program Year; SSP = Medicare Shared Savings Program; wBCF (+C) = weighted Bayesian Causal Forest (with expanded set of comparison practices).

Figure 5.I.6. Estimated impacts of CPC+ on Medicare outpatient ED visits in the evaluation sample in PY 5 (estimate and 90% CI) when using wBCF and DD models



CI = confidence interval or credible interval; DD = difference-in-differences; ED = emergency department; PY = Program Year; SSP = Medicare Shared Savings Program; wBCF = weighted Bayesian Causal Forest.

Figure 5.I.7. Estimated impacts of CPC+ on Medicare acute hospitalizations in the evaluation sample in PY 5 (estimate and 90 percent CI) when using wBCF and DD models



CI = confidence interval or credible interval; DD = difference-in-differences; PY = Program Year; SSP = Medicare Shared Savings Program; wBCF = weighted Bayesian Causal Forest.

C. Sensitivity Analysis C: Expanded comparison group for estimating impacts on Medicare expenditures

In this sensitivity analysis, we investigate BCF estimates' wide uncertainty intervals during the COVID-19 pandemic. Specifically, we found that for both evaluation sample and projected scaled sample estimates, wBCF and iBCF estimated greater uncertainty around all PY 5 and PY 4 impact estimates, respectively, than for PY 3 impact estimates; DD estimates' uncertainty did not correspondingly increase in PYs 4 and 5 compared with PY 3 (Table 5.I.8). Uncertainty intervals for PY 5 evaluation-sample impact estimates for Medicare expenditures were therefore almost twice as wide for wBCF compared with DD ($30 PBPM versus $16 PBPM in Track 1, $27 PBPM versus $18 PBPM in Track 2), with similar differences in PY 4 between iBCF and DD ($30 PBPM versus $15 PBPM in Track 1, $38 PBPM versus $27 PBPM in Track 2 SSP [non-SSP models did not converge]). In PY 3, uncertainty interval widths were on par between iBCF and DD ($12 PBPM versus $13 PBPM in Track 1, $13 PBPM versus $15 PBPM in Track 2).

To assess whether these wide credible intervals resulted from BCF's flexible modeling of COVID-19 (and other regional) control variables and the resulting collinearity issue discussed in Table 5.I.2, we included an expanded set of comparison practices from both CPC+ and comparison regions—we term this the "wBCF +C" sensitivity analysis. When including these additional practices, the estimated uncertainty around PY 5 impact estimates decreased to align more closely to that from DD (CI width of $16 for wBCF +C [versus $16 PBPM for DD and $30 PBPM for wBCF] in Track 1 and $18 for wBCF +C [versus $18 PBPM for DD and $27 PBPM for wBCF] in Track 2; Figure 5.I.5). *This wBCF +C analysis supports that regional differences in COVID-19 impact and response might be responsible for BCF being unable to distinguish between COVID-19 shocks and treatment effects, resulting in the large credible intervals.*

Between wBCF, wBCF +C, and the frequentist DDD analyses, evaluation sample impact estimates for SSP practices remained consistent, while impact estimates for non-SSP practices were sensitive to the different comparison group approaches, though in no model did non-SSP practices show savings (for both tracks, wBCF and DDD estimated increases in expenditures while BCF +C estimated null impacts).

Collectively, these analyses suggest that although Track 1 SSP practices show a moderately high probability of savings, there is less confidence that these practices offset CMFs—reinforcing scale-up findings that a cost-neutral scale-up is unlikely. Furthermore, these analyses suggest caution in Track 2 SSP practice savings (before CPC+ enhanced payments)—reinforcing the lack of a feasible scale-up in Track 2—and no analysis suggests savings amongst non-SSP practices.

5.I.4. Discussion

Consistent with previous year's findings, we estimated that there is almost no chance that a *nationwide scale-up* of either track's PY 5 effects would be cost neutral. We likewise did not find a promising *targeted scale-up* likely to offset CMFs using PY 5 data. A Track 1 scale-up to SSP practices, which showed the most promise in PY 4 with a 79 percent probability of being cost neutral, showed attenuated effects in PY 5: only a 34 percent probability of a cost-neutral scale-up. So although SSP practices were estimated to have an 80 percent probability of reducing gross expenditures in a Track 1 scale-up, in PY 5, no scale-up approach was predicted likely to generate net savings.

The estimated attenuation in PY 5 compared with PY 4 impacts could be true or spurious. For example, the following two mechanisms could explain true attenuation. First, the last year of the model could have

seen waning enthusiasm and engagement by practices. Compared with other program years, PY 5 saw higher practice and provider attrition (see Figure 2.1 of the main report). Second, although many beneficiaries avoided primary care at the onset of the COVID-19 pandemic, CPC+ practices' focus on continuity of care and investment in virtual health offerings could have translated to expenditure and ED visit reductions for their beneficiaries, relative to comparison practices with fewer virtual offerings. Once many beneficiaries resumed their in-person primary care after vaccines became available in early 2021 to those 65 and older, the differential impacts between beneficiaries served by CPC+ versus non-CPC+ practices might have diminished, for attenuated PY 5 impacts.

In contrast, the estimated attenuation could be spurious if iBCF's PY 4 impact estimates were overly optimistic or if wBCF's PY 5 estimates were overly pessimistic. On one hand, the estimated attenuation could be spurious due to residual bias. For example, BCF may have misattributed residual confounding due to COVID-19 to treatment effects, owing to the collinearity issue discussed in Sensitivity Analysis C. As shown earlier, DD and DDD analysis did not estimate such an impact attenuation between PY 4 and PY 5. Furthermore, iBCF's PY 4 point estimates were the most favorable across all analytic approaches (DD, DDD, iBCF) and years (PYs 3, 4, and 5), supporting that they might have been overly optimistic. On the other hand, the estimated attenuation could simply be due to noise, given that BCF's credible intervals widened greatly during the COVID period (PY 4 and PY 5). Within those wide uncertainty bounds, BCF's PY 4 and PY 5 estimates were consistent with one another. Thus, while changes in BCF's impact estimates for a Track 1 scale-up to SSP practices led to large changes in the *probability of offsetting CMFs*, the changes in *impact estimates* were not statistically meaningful.

Primary care changes take time to translate to meaningful impacts, and in later years of CPC+ (PY 4 and PY 5) a scale-up of Track 1 to SSP practices was estimated to reduce hospitalizations and outpatient ED visits with high probability (ED visit reductions emerged in PY 3 as well). In Track 2, a scale-up to SSP practices was also likely to reduce outpatient ED visits but not hospitalizations nor expenditures. Compared with Track 1, net savings were more difficult to achieve for Track 2 due to Track 2's much higher CMFs. Gross savings were, however, also estimated to be less likely for a Track 2 scale-up (47 percent) compared with a Track 1 scale-up (80 percent). Although Track 2 had more complex care delivery requirements and larger payments than Track 1, which might encourage larger impacts for Track 2 SSP practices and overall, CMS also required Track 2 practices to have more advanced care delivery at enrollment, which might have limited the room they had left for improvement, potentially precluding savings.

To estimate the impacts of scaling up CPC+, the scalability analysis addressed differences in characteristics between the evaluation sample and the projected scaled samples and captured uncertainty about which practices would volunteer for the scale-up. However, just as evaluation findings rely on assumptions necessary for estimating effects in the evaluation sample, these scalability conclusions rely on a further set of assumptions necessary for estimating effects in the projected scaled samples.

Sensitivity analyses assessing various components of these internal and external validity assumptions underlying scalability results affirmed that scaling up either track of CPC+ is unlikely to be cost-neutral. Namely, Track 1 and 2 nationwide and SSP-targeted scale-up estimates remained consistent across analytic approaches, under unmeasured effect modification as strong as the strongest measured effect modifier (SSP status), and when adjusting for region-specific shocks by including additional comparison practices from within CPC+ and comparison regions. Estimates for non-SSP practices were sensitive to different approaches for confounding adjustment, though no analytic approach estimated savings for non-SSP practices.

Several scale-up assumptions were not examined in sensitivity analyses. These include no relevant temporal differences between the scale-up and the evaluation, no differences in practice characteristics from what we saw in 2017 (particularly as they relate to targeting practices), no differences in participation drivers, no spillover effects, not retaining benefits after model termination, and no uncaptured changes due to COVID-19. As is always the case when impact analyses rely on assumptions, our results depend on these assumptions being true. Specifically, neither our scalability point estimates nor their credible intervals reflect any uncertainty about these assumptions; the assumptions are an additional source of uncertainty that our results do not capture. We now discuss these assumptions and their implications in turn.

Temporal changes between the evaluation and scale-up. The scalability analyses presented in this Appendix account for differences between the evaluation sample and scaled-up samples; they do not account for (1) potential changes under scale-up to model implementation or (2) health care changes since 2017. Changes to health information technology, care delivery, payment, lessons learned from the evaluation's implementation, differences in payer partnership, alignment and supports, and changes to learning support in the scale-up could potentially increase or decrease impacts for new participants compared with what we observed in evaluation sample practices with similar characteristics.

The health care environment has evolved since 2017, and will unquestionably continue to do so—for example, in the availability of competing care delivery models and programs. These alternative programs could moderate CPC+ impacts, as could changing state policies and other features of the context in which practices implemented CPC+ that might differ beyond 2022. The effects of these differences on model impacts could be large, but neither our point estimates nor their uncertainty intervals capture these factors. Together, model and context changes that would undoubtedly accompany a scale-up substantially undermine our confidence in the accuracy of our estimates. Instead, one should more precisely think of our estimates as providing an accurate retrospective estimate of the impact if CMS had offered CPC+ nationwide in 2017. In this sense, although our analysis sought to be rigorous in its approach to extrapolating geographically, it does not tackle the more difficult challenge of extrapolating impacts forward in time.

Practice characteristics and targeting. Our work assumes practice characteristics have largely remained unchanged since we collected baseline data on them in 2016 and 2017. We have accounted for the increase in EHR usage by assuming all practices would meet the Certified Electronic Health Record Technology eligibility criterion at the time of scale-up. However, if practice characteristics today differ from characteristics in 2017 across effect modifiers such as SSP status, effect estimates at the time of scale-up will likewise differ. For example, the SSP program has changed since 2017. With more downside risks, the types of practices that participated in SSP as of January 1, 2017, and the types of practices participating today might have likewise changed. Between 2017 and 2021, among CPC+ practices, almost one-third of SSP practices dropped out of SSP and more than one-quarter of non-SSP practices joined the program. For 2017 SSP practice impacts to correspond to 2023 SSP practice impacts, we have to assume the evaluation-sample impacts relate to 2017 characteristics in the same way that scale-up impacts will relate to 2023 characteristics. Our analysis did not capture uncertainty around this assumption. However, given the observed heterogeneity in practice impacts in PY 5, we have no reason to believe another subgroup of practices, including subgroups by 2023 SSP status, would likely offset CMFs either.

Drivers of participation. Because CMS chose CPC+ regions for their payer alignment, new regions might see different drivers of practice participation. For example, lower payer alignment in scale-up

regions could lead to lower practice participation rates. The changing policy landscape, the COVID-19 pandemic, and any alternative models available such as Primary Care First and ACO REACH, might also differentially drive participation in a scaled CPC+ model. Although lower or higher overall CPC+ participation would not affect projected PBPM impacts, different drivers of participation would lead to uncertainty in scale-up effects that our analyses do not capture.

Spillover effects. The model could affect nonparticipants in regions with CPC+ practices because of spillover effects. For example, large health systems make changes for CPC+ that benefit all practices, including those that are not in CPC+. These spillover effects could increase impacts (from nonvolunteering practices that benefit) or decrease impacts (from non-CPC+ practices in CPC+ regions that have already benefited from the evaluation's model).

Retention of changes after model termination. Because some CPC+ changes required fixed rather than variable costs, evaluation participants might have sustained some of their changes under the model when CMS discontinued it. In this case, our estimates would overstate true scale-up impacts, because practices would retain the benefits of fixed-cost changes after the evaluation period ended even without a scale-up. These fixed-cost changes include changes to discharge protocols (such as coordination and communication with specialists), collecting data more systematically for electronic clinical quality measures, and process of care and workflow improvements (such as improved appointment scheduling processes and same-day appointment availability). However, the CPC+ components theorized to drive the most changes require continued investments: hiring new staff for enhanced care management (e.g., nurse care managers); behavioral health integration (e.g., embedding behavioral health staff); and episodic care management (e.g., following up with patients after hospital or ED visits to avoid readmissions and exacerbation). In interviews in the final year of the model, practices were uncertain whether they would be able to retain changes that require continued investment once CMS terminated the model.

COVID-19. This analysis adjusted for confounding and effect modification of COVID-19, but measured factors imperfectly capture the upheaval caused by the pandemic. The pandemic affects service use, care delivery, drivers of participation, and patient and practice characteristics we cannot fully account for with our data. For example, the pandemic might have increased (or decreased) ED visits, hospitalizations, and spending due to outbreaks beyond what excess mortality, the pandemic vulnerability index, government response index, and social vulnerability index can control for. Sensitivity analyses using an expanded comparison group (wBCF +C) and DDD results reinforce the internal validity of evaluation impact findings, and by extension scalability results. With respect to external validity, if impact heterogeneity because of COVID-19 were as strong as the strongest measured effect modifier, then, as Sensitivity Analysis A shows, results would remain consistent with the main scalability findings.

The impacts of a primary care intervention on Medicare expenditures take time to manifest, and although no CPC+ scale-up is likely to offset CMFs, current and future models such as Primary Care First will continue to build on quality improvements begun under CPC+. These current and future models could also consider similar analyses to guide an evidence-based targeted scale-up. Future work can also consider scaling to practice types and geographic regions estimated to improve quality of care, because the model expansion criteria from the Center for Medicare and Medicaid Innovation (Innovation Center) are Medicare savings with no change in quality or Medicare cost neutrality with quality improvements. This Appendix describes potential paths forward and can inform future directions for the Innovation Center's primary care models.

5.J. Impact of CPC+ on acute hospitalizations

In this Appendix, we examine the impact of the Comprehensive Primary Care Plus (CPC+) model on types of acute hospital admissions for Medicare fee-for-service (FFS) beneficiaries during the five years of CPC+. In Section 5.J.1, we describe the motivation for this analysis, including an overview of how CPC+ could affect types of hospitalizations, and present key findings. We then explain the analytic methods, study population, and outcomes of interest (Section 5.J.2). Next, we describe the results (Section 5.J.3) and discuss their implications (Section 5.J.4). Finally, we discuss the limitations of this analysis (5.J.5) and then conclude (5.J.6).

5.J.1. Introduction

Spending for hospital care services in the United States surpassed $1 trillion in 2016, representing 32 percent of total health care spending (Hartman et al. 2018; AMA 2022b). More complex and higher acuity admissions account for a substantial proportion of these costs. Surgical stays, for example, account for half of all Medicare inpatient spending despite making up one-third of hospitalizations (Liang et al. 2017; Chhabra et al. 2019).

With growing spending on hospital care, many initiatives are focusing on acute hospitalizations. For example, reducing acute care admissions is an explicit goal of Primary Care First (PCF), a large, national primary care model the Centers for Medicare & Medicaid Services (CMS) launched in 2021. The premise behind this and related primary care initiatives is that high-quality outpatient care should reduce unnecessary hospitalizations for conditions thought amenable to improved primary care.

The CPC+ model—the predecessor to PCF—had a similar goal. By improving care delivery across five Comprehensive Primary Care Functions, CMS hypothesized that primary care practices could reduce acute hospitalizations. We identified two functions in particular—(1) Access and Continuity and (2) Care Management—that could contribute to reduced hospitalizations by improving timely access for acute episodes, strengthening case management to avoid acute exacerbations of chronic conditions, and enhancing coordination efforts to follow up after discharge to reduce inpatient readmissions. As in Chapter 5, our findings indicate CPC+ practices reduced all-cause acute hospitalizations by 0.9 percent in Track 1 and 1 percent in Track 2, and reduced acute inpatient expenditures in Track 1 by 1.1 percent, relative to a comparison group of practices.

The analysis in this Appendix extends the main CPC+ impact analysis with a fuller picture of the impact CPC+ had on acute hospitalizations. Specifically, we study which types of acute hospital admissions were most affected over the model's five years; identify which practices were most successful in reducing hospitalization; and test whether impacts differed by beneficiary type. We hypothesized the greatest changes from CPC+ would occur in medical (nonsurgical) admissions and in lower severity admissions—such as those for conditions with symptoms and exacerbations that timely advanced primary care might mitigate. We also hypothesized that some types of practices, such as independent primary care practices, would be better able to reduce hospitalizations than others, such as hospital- or system-owned practices, because of differences in administration and financial incentives. Further, we hypothesized that improved primary care might have more capacity to reduce hospitalizations among Medicare beneficiaries with poorer health—for example, those with higher hierarchical condition category (HCC) scores, who are likely to have more encounters with hospitals. A better understanding of exactly which types of hospitalizations CPC+ affected and which type of practices and Medicare beneficiaries were significant in driving these results could be particularly important in the context of PCF and future primary care

initiatives. The findings in this Appendix might also shed light on why CPC+ did not make meaningful reductions in total Medicare expenditures, despite reducing hospital admissions.

Key findings

- Relative to Medicare fee-for-service beneficiaries in comparison practices, beneficiaries attributed to CPC+ practices experienced:
  - Greater reductions in acute medical hospital admissions, with estimated average annual reductions of 1.3 and 1.4 percent in Tracks 1 and 2, respectively, and no effects on acute surgical admissions
  - Slower growth in expenditures for acute medical hospital admissions, with estimated reductions of 2.2 percent in Track 1 and 2.6 percent in Track 2, and no effects on expenditures for acute surgical admissions
- The reductions in medical hospital admissions were largest in PY 3 and PY 4 and accounted for nearly all reductions in all-cause acute hospitalizations.
- Reductions in the lowest severity of admissions (that is, those without any complication or comorbidity) drove the reductions in acute medical admissions. However, all severity levels contributed to reduced expenditures for acute medical admissions.
- There were some differences by Medicare Shared Savings Program (SSP) status: non-SSP practices reduced the least complex hospitalizations (medical admissions without a complication or comorbidity), and there is some evidence SSP practices reduced more complex admissions.
- There were some differences by practice subgroup. We consistently found impacts were concentrated among independent practices, versus hospital- or system-owned practices, which showed no effects.
- Reductions in expenditures for acute medical admissions were largest for beneficiaries in the highest quartile of the hierarchical condition category score distribution in both tracks. For Track 2, impacts were concentrated among patients who were dually eligible for Medicare and Medicaid.

5.J.2. Methods

A. Evaluation design

For this analysis, we used the same evaluation approach as in the main CPC+ impact analysis and studied the same study population (Chapter 5). Specifically, we studied the 1,373 Track 1 and 1,515 Track 2 practices that joined CPC+ in 2017 and remained in the model for the first 90 days. Every quarter, we attributed beneficiaries to the practice that delivered the largest share of their primary care visits over the prior two years (Appendix 5.B). We then assigned beneficiaries to the CPC+ or comparison groups at two points in time. For the baseline period, we assigned beneficiaries to the first practice to which they were attributed during the baseline period. We followed an intent-to-treat rule by continuing to assign the beneficiary to the same practice throughout the baseline period regardless of whether the beneficiary continued to receive care at that practice. We repeated the same process for the intervention period, assigning patients to the first practice to which they were attributed after the intervention began (Appendix 5.B). We relied on the same external comparison group used for the main impact analysis,

which showed similarity between CPC+ and comparison practices across practice site, market, and patient characteristics before CPC+ began (Ghosh et al. 2020; Kranker et al. 2020).

B. Data and study population

To construct outcomes, we analyzed Medicare claims from the CMS Virtual Research Data Center over the baseline period (2016) and the five program years of CPC+ (January 2017 to December 2021). Our final sample consisted of 1,549,585 and 1,896,880 Medicare FFS beneficiaries attributed to 1,373 and 1,515 Track 1 and Track 2 CPC+ practices, respectively, and 5,347,499 and 4,507,499 beneficiaries attributed to 5,243 and 3,783 comparison practices in Tracks 1 and 2, respectively.

C. Hospitalization measures

We focused on hospitalizations at short-stay acute hospitals and critical access hospitals. We categorized hospital admissions using information from the inpatient prospective payment system (IPPS), which uses Medicare severity diagnosis-related groups (MS-DRGs) to determine hospital payments. Specifically, CMS uses MS-DRGs to determine the amount hospitals will be reimbursed based on expected resource usage. Each admission is assigned to one of approximately 770 MS-DRGs, and the hospital is paid a fixed amount that varies by MS-DRG. The mapping between principal diagnoses and MS-DRGs is not one to one: MS-DRG assignment is based on the combination of diagnoses, procedures, age, sex, discharge status, and the presence of major or minor complications or comorbidities.

We obtained details on the list of MS-DRGs from Table 5 on the IPPS Final Rule page for each year from 2016 to 2022[119] and merged the information with Medicare claims data using the MS-DRG. We then created two groupings of hospitalizations:

1. **Surgical versus medical admissions.** The MS-DRG system identifies whether an admission is assigned to a surgical or a medical (nonsurgical) MS-DRG. For example, coronary bypass is a surgical MS-DRG, and urinary tract infection is a medical MS-DRG. A surgical MS-DRG, on average, will be higher acuity and require more hospital resources than a medical MS-DRG. It is therefore more costly, and hospitals receive higher payment.
2. **Severity based on complication or comorbidity.** MS-DRGs are organized into families (e.g., "respiratory infections and inflammations") with two or three levels: the base MS-DRG and one or two higher-complexity MS-DRGs. Assignment to the higher-complexity MS-DRGs occurs if complications or comorbidities (CCs) or major complications or comorbidities (MCCs) are present. Hospital reimbursements for MS-DRGs with CCs or MCCs are often substantially greater than for the base MS-DRG (which is without a complication or comorbidity). For example, for the MS-DRG triplet of "Respiratory infections and inflammations," the MS-DRG with an MCC receives a reimbursement that is more than 1.5 times the MS-DRG with a CC and that is more than double the base MS-DRG.

The full list of outcomes we studied are listed in Table 5.J.1. We first analyzed all-cause acute hospitalizations (all types of acute hospitalizations together). Second, we examined surgical versus

[119] Although we include acute hospitalizations with admission dates through 2021, we use DRG information through 2022. Some hospitalizations in our analysis, such as those with admission dates at the end of 2021, will have discharge dates in 2022, and when assigning DRG information, we use the latest claim on the admission which can consequently include dates in 2022.

medical hospitalizations. Third, we studied the combination of surgical/medical hospitalizations by severity type (for example, "medical acute hospitalizations without an MCC or CC"). For all measures, we accounted for beneficiaries' Medicare FFS eligibility throughout the program year. We express expenditure measures as dollars per beneficiary per month (PBPM). Service use measures are the annualized number of hospitalizations per 1,000 beneficiaries.

Table 5.J.1. Description of acute inpatient hospital outcome variables studied

Outcome
Expenditures (per beneficiary per month)
Medicare expenditures for all-cause acute hospital admissions
Medicare expenditures for acute medical hospital admissions
Medicare expenditures for acute surgical hospital admissions
Medicare expenditures for acute surgical hospital admissions with an MCC
Medicare expenditures for acute surgical hospital admissions with a CC
Medicare expenditures for acute surgical hospital admissions without an MCC or CC
Medicare expenditures for acute medical hospital admissions with an MCC
Medicare expenditures for acute medical hospital admissions with a CC
Medicare expenditures for acute medical hospital admissions without an MCC or CC
Annualized service use (per 1,000 beneficiaries per year)
Annualized number of all-cause acute hospitalizations
Annualized number of acute hospitalizations that are surgical
Annualized number of acute hospitalizations that are medical
Annualized number of acute surgical hospitalizations with an MCC
Annualized number of acute surgical hospitalizations with a CC
Annualized number of acute surgical hospitalizations without an MCC or CC
Annualized number of acute medical hospitalizations with an MCC
Annualized number of acute medical hospitalizations with a CC
Annualized number of acute medical hospitalizations without an MCC or CC

CC = complication or comorbidity; MCC = major complication or comorbidity.

D. Statistical analysis

Following the main CPC+ impact analysis, we used a difference-in-differences (DD) framework and compared the changes in mean acute hospitalizations and expenditures on hospitalizations for Medicare beneficiaries in CPC+ practices between the 12 months before CPC+ (baseline) and each of the program years of CPC+ with changes among beneficiaries in the comparison practices over the same period. Because practice transformation takes time, we expected that effects of CPC+ would emerge and potentially grow in later program years. We also estimated the average annual impact of CPC+ across the five program years.

We estimated DD models separately by track, reflecting the different eligibility requirements, practice care delivery requirements, and financial payments and incentives. To net out prior observable differences between CPC+ and comparison beneficiaries not fully eliminated by matching, the regression models controlled for practice fixed effects and beneficiaries' characteristics at baseline, such as demographics; the original reason for Medicare eligibility; the HCC score, a comprehensive summary of demographic and clinical factors measuring risk for subsequent Medicare expenditures; and multiple individual chronic conditions. To account for the potential for the coronavirus disease 2019 (COVID-19) pandemic to confound our estimates in the last two program years, we also included controls to account for regional differences in COVID-19. (Appendix 5.E lists the full set of control variables used in the main CPC+ impact analysis, which we followed for this analysis.) *P*-values were two-sided and considered statistically significant at $P < 0.1$.

E. Subgroup analyses

We estimated impacts for each track overall, then estimated impacts separately by practices' baseline Medicare Shared Savings Program (SSP) status to investigate whether participating in both CPC+ and an SSP Accountable Care Organization had a different impact than participating in CPC+ alone. Given that SSP participation is a critical dimension on which participating CPC+ practices differ, we estimated these separate regressions, by SSP status, for all outcomes.

In addition, the impacts of CPC+ on acute hospitalizations could differ for various types of practices and beneficiaries, based on other baseline characteristics. Therefore, we estimated impacts for various types of practices, such as those that had a larger number of primary care practitioners or had participated in prior primary care transformation initiatives at baseline, or by ownership status (hospital- or system- owned versus independent practices). In addition, we estimated the effects of CPC+ on subsets of beneficiaries for whom CPC+ is likely to have especially large effects, such as the chronically ill and other patients with complex health conditions.

We studied the same set of practice and beneficiary subgroups as in the main CPC+ impact analysis. To derive impacts for each subgroup, we used the same estimation approach (see Appendix 5.E for more details). In our subgroup analyses of practices and beneficiaries, we focused on broad surgical and medical hospitalizations (rates and expenditures). That is, we did not study the detailed severity categories within these types of hospital admissions to keep the results tractable. For a similar reason, we focused on the average annual impacts for PY 1 through PY 5 rather than individual program years for our subgroup analyses.

5.J.3. Results over the five program years

A. All-cause hospitalizations

Across the five years of the model, CPC+ reduced all-cause acute hospitalizations in both tracks (0.9 percent in Track 1 and 1 percent in Track 2), and reduced expenditures in Track 1 ($3.1 PBPM or 1.1 percent). The effects were concentrated in later years of the model, specifically PY 3 and PY 4. There were also reductions in acute hospital expenditures for Track 2, but the results were not statistically significant. When stratifying by SSP status, there were reductions in the average annual rate of all-cause acute hospitalizations for SSP practices in Track 1 (1.6 percent) and for non-SSP practices in Track 2 (1.3 percent). However, there were no statistical effects on the rate of hospitalizations for Track 1 non-SSP practices, nor for Track 2 SSP practices. In both tracks, CPC+ reduced expenditures for acute hospitalizations among SSP practices, with reductions of $6.8 PBPM for Track 1 and $6.0 PBPM for Track 2 (Tables 5.J.2 and 5.J.3, and Chapter 5 of the fifth annual report).

B. Impact on surgical and medical hospitalizations

In both tracks, reductions in medical admissions drove reductions in all-cause acute hospitalizations from CPC+, with no effects on surgical admissions. Over the course of CPC+, acute medical hospitalizations declined for CPC+ and comparison practices relative to the year before CPC+ began. However, this decline was greater for CPC+ practices than for comparison practices, leading to annualized average reductions of 2.4 and 2.8 medical hospitalizations per 1,000 beneficiaries in Tracks 1 and 2, respectively, which translates to reductions of 1.3 and 1.4 percent ($p = 0.04$ for Track 1 and $p = 0.03$ Track 2). The reductions in acute medical hospitalizations were concentrated in PY 3 and PY 4 for both tracks, with the largest reductions being 2.3 percent ($p < 0.01$) in both tracks, in PY 4 for Track 1

and in PY 3 for Track 2. Although acute surgical hospitalizations declined for both CPC+ and comparison practices, the magnitude of these declines were similar for CPC+ and comparison practice groups, resulting in CPC+ having no effect on surgical admissions in either track (Tables 5.J.2 and 5.J.3).

Reduced expenditures for medical admissions in both tracks drove the reductions in expenditures for all-cause hospital admissions, and there were no reductions in expenditures for surgical admissions. Over the five years of CPC+, there was an increase in expenditures for acute medical hospitalizations for both CPC+ and comparison practices relative to the year before CPC+ began. However, this increase was greater for comparison practices than for CPC+ practices. As a result, CPC+ reduced annual average reduction in expenditures for medical admissions, relative to the comparison group, by $3.2 PBPM (2.2 percent) in Track 1 and $3.8 PBPM (2.6 percent) in Track 2. These reductions occurred multiple years; the largest impacts were in PY 4 for Track 1 ($5.9 PBPM, 4.0 percent) and PY 3 for Track 2 ($6.3 PBPM, 4.2 percent). For both tracks, there were no statistical reductions in expenditures for acute surgical admissions in any program year (Tables 5.J.2 and 5.J.3).

Our results indicate that although medical admissions represented 70 percent of all-cause acute hospitalizations before CPC+, the reduction in medical hospitalizations accounted for nearly all the reduction in all-cause acute hospitalizations. For example, CPC+ led to a reduction of 4.8 admissions per 1,000 beneficiaries for all acute hospitalizations in the third program year for Track 2 practices, and all 4.8 of those admissions came from reduced medical admissions (Table 5.J.3). Similarly, slower growth in medical expenditures drove the effect on all-cause acute expenditures. For example, CPC+ reduced acute expenditures by $6.9 PBPM in PY 4 for Track 1, and $5.9 of that was from reduced medical expenditures (Table 5.J.2). That is, medical admissions accounted for about 86 percent of the reduction in expenditures for all-cause acute hospitalizations, despite representing 46 percent of overall acute inpatient expenditures at baseline.

When stratifying by SSP status, CPC+ reduced acute medical hospitalizations for SSP practices in Track 1 and for non-SSP practices in Track 2. Over the five years, CPC+ reduced average annual medical hospitalizations by 1.9 percent in Track 1 SSP practices and by 2.3 percent for Track 2 non-SSP practices (Tables 5.J.2 and 5.J.3). There were reductions multiple years for both sets of practices, with the largest in PY 4 for Track 1 SSP practices (3.4 percent; $p < 0.01$) and in PY 3 for Track 2 non-SSP practices (3.5 percent; $p < 0.01$). The DD estimates also show reductions for acute medical hospitalizations in the other two practice groups (non-SSP in Track 1 and SSP in Track 2). However, these estimates were not statistically significant.

The average annual DD estimate for surgical admissions was not statistically significant for any practice group in either track. However, CPC+ reduced surgical admissions among SSP practices in later years of the model. Relative to the comparison group, SSP practices in both tracks had a reduction in surgical admissions in PY 4 (2.7 percent in Track 1, $p = 0.02$; and 2.4 percent in Track 2, $p = 0.08$) and SSP practices in Track 1 also had a relative reduction in PY 5 (3.2 percent; $p = 0.01$; Tables 5.J.2 and 5.J.3). For Track 1 SSP practices, the magnitude of the reductions in surgical admissions was smaller than the reductions in acute medical admissions each year, both in the number of admissions per 1,000 beneficiaries per year and the percentage relative to the baseline rate.

Across the five program years, there were reductions in expenditures for acute medical admissions for SSP practices in both tracks and for non-SSP practices in Track 2. SSP practices in Track 1 had annual average reductions in expenditures for acute medical admissions of 3.1 percent ($p < 0.01$; Table 5.J.2); SSP Track 2 practices had reductions of 2.5 percent ($p = 0.051$); and non-SSP Track 2 practices

had reductions of 2.7 percent ($p < 0.01$; Table 5.J.3). There were reductions in multiple program years, with the largest effects in PY 3 and PY 4. For non-SSP practices in Track 1, when there was not a statistical reduction in average annual acute medical expenditures, there was a reduction in PY 4 of 2.6 percent ($p = 0.08$). Estimates for other years were also favorable but not statistically significant (Table 5.J.3).

Over the five years of CPC+, there were no reductions in the annual average expenditures on surgical admissions. However, SSP practices in Track 1 had reductions in expenditures for surgical admissions for PY 4 and PY 5, which aligns with reductions in surgical admissions these years (Table 5.J.2). The DD estimates show that SSP practices in Track 2 also had reductions in surgical expenditures in later years, but these results were not statistically significant (Table 5.J.3). One unexpected result is that Track 2 non-SSP practices had an *increase* in expenditures for surgical admissions, despite no change in the number of surgical admissions. Comparison practices reduced expenditures more than CPC+ practices between baseline and PY 5, resulting in an average annual DD estimate of $5.1 PBPM (3.6 percent, $p < 0.01$).

Table 5.J.2. Regression-adjusted means and estimated impact of CPC+ on acute surgical and medical hospitalizations for attributed Medicare FFS beneficiaries over the five program years, Track 1

	Track 1—Overall						Track 1—SSP						Track 1—Non-SSP					
	CPC+ mean[a]	C mean[a]	Impact estimate[b] (SE)	Percentage impact[c]	90 percent confidence interval	*p*-Value	CPC+ mean[a]	C mean[a]	Impact estimate[b] (SE)	Percentage impact[c]	90 percent confidence interval	*p*-Value	CPC+ mean[a]	C mean[a]	Impact estimate[b] (SE)	Percentage impact[c]	90 percent confidence interval	*p*-Value
Monthly Medicare expenditures by service category (per beneficiary per month)																		
Total expenditures for acute hospital admissions																		
Baseline	$275	$282	NA	NA	NA	NA	$282	$285	NA	NA	NA	NA	$268	$278	NA	NA	NA	NA
PY 1	$279	$285	$1.2 ($2.0)	0.4%	(-$2.1, $4.5)	0.55	$285	$290	-$1.4 ($2.6)	-0.5%	(-$5.7, $2.9)	0.59	$273	$279	$4.0 ($3.1)	1.5%	(-$1.1, $9.2)	0.20
PY 2	$285	$293	-$1.5 ($2.0)	-0.5%	(-$4.9, $1.8)	0.44	$292	$298	-$2.6 ($2.7)	-0.9%	(-$7.0, $1.9)	0.34	$276	$287	-$0.4 ($3.0)	-0.2%	(-$5.4, $4.5)	0.89
PY 3	$295	$306	-$4.4** ($2.2)	-1.5%	(-$8.0, -$0.8)	0.05	$302	$314	-$8.1*** ($3.0)	-2.6%	(-$13.1, -$3.2)	0.01	$287	$297	-$0.1 ($3.2)	0.0%	(-$5.4, $5.2)	0.97
PY 4	$280	$293	-$6.9*** ($2.3)	-2.4%	(-$10.7, -$3.1)	0.00	$286	$302	-$12.3*** ($3.2)	-4.1%	(-$17.6, -$6.9)	0.00	$273	$284	-$1.1 ($3.3)	-0.4%	(-$6.5, $4.4)	0.75
PY 5	$294	$305	-$3.8 ($2.4)	-1.3%	(-$7.8, $0.2)	0.12	$306	$320	-$10.5*** ($3.5)	-3.3%	(-$16.2, -$4.8)	0.00	$283	$291	$1.3 ($3.5)	0.5%	(-$4.5, $7.0)	0.71
PY 1 through 5	$287	$296	-$3.1* ($1.8)	-1.1%	(-$6.0, -$0.2)	0.08	$295	$305	-$6.8*** ($2.4)	-2.3%	(-$10.7, -$2.9)	0.00	$278	$288	$0.7 ($2.7)	0.3%	(-$3.7, $5.1)	0.79
Expenditures for acute surgical hospital admissions																		
Baseline	$148	$149	NA	NA	NA	NA	$152	$151	NA	NA	NA	NA	$143	$146	NA	NA	NA	NA
PY 1	$148	$147	$2.6* ($1.5)	1.8%	($0.1, $5.1)	0.09	$151	$150	$0.4 ($2.1)	0.3%	(-$3.0, $3.8)	0.85	$145	$143	$5.0** ($2.3)	3.6%	($1.2, $8.8)	0.03
PY 2	$149	$149	$0.4 ($1.5)	0.3%	(-$2.0, $2.8)	0.79	$153	$153	-$0.3 ($2.0)	-0.2%	(-$3.6, $3.0)	0.88	$144	$145	$1.2 ($2.1)	0.8%	(-$2.3, $4.7)	0.58
PY 3	$154	$155	-$0.5 ($1.6)	-0.3%	(-$3.1, $2.1)	0.74	$158	$159	-$2.4 ($2.2)	-1.5%	(-$6.0, $1.2)	0.28	$150	$151	$1.5 ($2.2)	1.0%	(-$2.2, $5.2)	0.49
PY 4	$138	$140	-$1.0 ($1.6)	-0.7%	(-$3.6, $1.5)	0.51	$141	$144	-$4.0* ($2.2)	-2.8%	(-$7.6, -$0.4)	0.07	$136	$136	$2.5 ($2.2)	1.9%	(-$1.1, $6.2)	0.25
PY 5	$143	$145	-$1.0 ($1.7)	-0.7%	(-$3.7, $1.8)	0.56	$147	$151	-$4.8** ($2.4)	-3.2%	(-$8.7, -$0.9)	0.04	$138	$139	$2.0 ($2.4)	1.5%	(-$1.9, $5.9)	0.40
PY 1 through 5	$146	$147	$0.1 ($1.2)	0.1%	(-$2.0, $2.1)	0.94	$150	$151	-$2.1 ($1.7)	-1.4%	(-$4.9, $0.7)	0.21	$142	$143	$2.4 ($1.8)	1.7%	(-$0.6, $5.4)	0.19

Table 5.J.2. ***(continued)***

	Track 1—Overall						Track 1—SSP						Track 1—Non-SSP					
	CPC+ mean[a]	C mean[a]	Impact estimate[b] (SE)	Percentage impact[c]	90 percent confidence interval	*p*-Value	CPC+ mean[a]	C mean[a]	Impact estimate[b] (SE)	Percentage impact[c]	90 percent confidence interval	*p*-Value	CPC+ mean[a]	C mean[a]	Impact estimate[b] (SE)	Percentage impact[c]	90 percent confidence interval	*p*-Value
Expenditures for acute medical hospital admissions																		
Baseline	$127	$133	NA	NA	NA	NA	$130	$134	NA	NA	NA	NA	$125	$132	NA	NA	NA	NA
PY 1	$131	$138	-$1.4 ($1.0)	-1.1%	(-$3.1, $0.3)	0.17	$134	$140	-$1.8 ($1.3)	-1.3%	(-$3.9, $0.3)	0.17	$128	$136	-$1.0 ($1.6)	-0.7%	(-$3.5, $1.6)	0.54
PY 2	$136	$144	-$1.9* ($1.1)	-1.4%	(-$3.8, -$0.1)	0.08	$139	$146	-$2.3 ($1.5)	-1.6%	(-$4.7, $0.2)	0.13	$132	$141	-$1.6 ($1.6)	-1.2%	(-$4.3, $1.1)	0.33
PY 3	$141	$151	-$3.9*** ($1.2)	-2.7%	(-$5.9, -$1.8)	0.00	$144	$155	-$5.8*** ($1.7)	-3.8%	(-$8.5, -$3.0)	0.00	$137	$146	-$1.7 ($1.8)	-1.2%	(-$4.6, $1.3)	0.36
PY 4	$141	$153	-$5.9*** ($1.4)	-4.0%	(-$8.1, -$3.6)	0.00	$146	$158	-$8.3*** ($1.9)	-5.4%	(-$11.3, -$5.2)	0.00	$137	$148	-$3.6* ($2.1)	-2.6%	(-$7.0, -$0.2)	0.08
PY 5	$151	$160	-$2.8* ($1.4)	-1.8%	(-$5.2, -$0.4)	0.05	$158	$168	-$5.7*** ($2.1)	-3.5%	(-$9.1, -$2.3)	0.01	$144	$152	-$0.7 ($2.1)	-0.5%	(-$4.1, $2.7)	0.73
PY 1 through 5	$140	$149	-$3.2*** ($1.0)	-2.2%	(-$4.8, -$1.5)	0.00	$145	$154	-$4.7*** ($1.3)	-3.1%	(-$6.9, -$2.5)	0.00	$136	$145	-$1.7 ($1.5)	-1.2%	(-$4.1, $0.7)	0.25
Annualized service use (per 1,000 beneficiaries per year)																		
Total annualized acute hospital admissions																		
Baseline	290	289	NA	NA	NA	NA	291	289	NA	NA	NA	NA	289	288	NA	NA	NA	NA
PY 1	289	288	-0.6 (1.5)	-0.2%	(-3.1, 1.9)	0.68	289	290	-2.7 (1.9)	-0.9%	(-5.8, 0.5)	0.16	289	286	1.6 (2.4)	0.6%	(-2.3, 5.5)	0.51
PY 2	285	285	-1.8 (1.6)	-0.6%	(-4.5, 0.9)	0.27	286	287	-2.2 (2.1)	-0.8%	(-5.7, 1.3)	0.30	283	283	-1.4 (2.6)	-0.5%	(-5.6, 2.8)	0.57
PY 3	284	286	-2.6 (1.8)	-0.9%	(-5.5, 0.3)	0.14	286	289	-4.9** (2.2)	-1.7%	(-8.6, -1.2)	0.03	283	282	0.0 (2.8)	0.0%	(-4.6, 4.5)	1.00
PY 4	243	247	-4.9*** (1.8)	-2.0%	(-7.8, -2.0)	0.01	245	251	-8.0*** (2.3)	-3.2%	(-11.9, -4.2)	0.00	241	242	-1.4 (2.8)	-0.6%	(-5.9, 3.2)	0.62
PY 5	244	246	-2.6 (1.8)	-1.1%	(-5.6, 0.4)	0.15	250	253	-5.1** (2.5)	-2.0%	(-9.2, -1.0)	0.04	239	238	0.7 (2.7)	0.3%	(-3.7, 5.0)	0.81
PY 1 through 5	268	269	-2.5* (1.4)	-0.9%	(-4.9, -0.1)	0.08	270	273	-4.5** (1.8)	-1.6%	(-7.5, -1.5)	0.01	266	265	-0.1 (2.2)	0.0%	(-3.8, 3.6)	0.96

Table 5.J.2. ***(continued)***

	Track 1—Overall						Track 1—SSP						Track 1—Non-SSP					
	CPC+ mean[a]	C mean[a]	Impact estimate[b] (SE)	Percentage impact[c]	90 percent confidence interval	*p*-Value	CPC+ mean[a]	C mean[a]	Impact estimate[b] (SE)	Percentage impact[c]	90 percent confidence interval	*p*-Value	CPC+ mean[a]	C mean[a]	Impact estimate[b] (SE)	Percentage impact[c]	90 percent confidence interval	*p*-Value
Annualized acute surgical hospital admissions																		
Baseline	90	88	NA	NA	NA	NA	90	88	NA	NA	NA	NA	90	87	NA	NA	NA	NA
PY 1	89	85	1.0 (0.7)	1.2%	(-0.1, 2.1)	0.12	89	86	0.1 (0.8)	0.1%	(-1.3, 1.5)	0.91	89	84	2.0* (1.1)	2.4%	(0.3, 3.8)	0.05
PY 2	86	84	0.1 (0.7)	0.2%	(-0.9, 1.2)	0.82	87	84	0.2 (0.9)	0.2%	(-1.2, 1.6)	0.85	86	83	0.1 (1.0)	0.1%	(-1.6, 1.8)	0.91
PY 3	87	85	-0.1 (0.7)	-0.1%	(-1.2, 1.1)	0.92	87	85	-0.6 (0.9)	-0.6%	(-2.0, 0.9)	0.53	87	84	0.5 (1.1)	0.6%	(-1.3, 2.2)	0.65
PY 4	71	70	-0.9 (0.7)	-1.2%	(-2.0, 0.2)	0.18	71	71	-2.0** (0.8)	-2.7%	(-3.4, -0.6)	0.02	72	69	0.3 (1.0)	0.4%	(-1.4, 2.0)	0.81
PY 5	69	67	-0.6 (0.7)	-0.9%	(-1.8, 0.5)	0.35	70	70	-2.3** (0.9)	-3.2%	(-3.8, -0.8)	0.01	69	65	0.8 (1.0)	1.2%	(-0.9, 2.6)	0.42
PY 1 through 5	80	78	-0.1 (0.5)	-0.1%	(-1.0, 0.8)	0.88	80	79	-0.9 (0.7)	-1.0%	(-2.0, 0.3)	0.22	80	77	0.7 (0.9)	0.9%	(-0.7, 2.1)	0.40
Annualized acute medical hospital admissions																		
Baseline	200	201	NA	NA	NA	NA	200	201	NA	NA	NA	NA	199	201	NA	NA	NA	NA
PY 1	200	203	-1.7 (1.2)	-0.8%	(-3.7, 0.4)	0.18	200	204	-2.8* (1.6)	-1.4%	(-5.4, -0.1)	0.09	200	202	-0.5 (1.9)	-0.2%	(-3.6, 2.7)	0.81
PY 2	198	201	-2.0 (1.4)	-1.0%	(-4.3, 0.3)	0.15	199	203	-2.4 (1.8)	-1.2%	(-5.4, 0.6)	0.19	197	200	-1.6 (2.1)	-0.8%	(-5.0, 1.9)	0.46
PY 3	197	201	-2.6* (1.5)	-1.3%	(-5.0, -0.1)	0.08	199	204	-4.3** (1.9)	-2.1%	(-7.5, -1.2)	0.02	196	198	-0.5 (2.3)	-0.2%	(-4.2, 3.3)	0.83
PY 4	172	177	-4.0*** (1.5)	-2.3%	(-6.5, -1.5)	0.01	174	181	-6.1*** (2.0)	-3.4%	(-9.4, -2.7)	0.00	170	173	-1.6 (2.3)	-0.9%	(-5.4, 2.2)	0.48
PY 5	175	178	-2.0 (1.5)	-1.1%	(-4.5, 0.6)	0.20	180	183	-2.8 (2.1)	-1.5%	(-6.3, 0.7)	0.19	171	172	-0.2 (2.2)	-0.1%	(-3.8, 3.5)	0.93
PY 1 through 5	188	191	-2.4** (1.2)	-1.3%	(-4.4, -0.4)	0.04	190	194	-3.7** (1.6)	-1.9%	(-6.3, -1.1)	0.02	186	188	-0.9 (1.8)	-0.5%	(-3.9, 2.2)	0.64
Unweighted sample sizes																		
Number of practices	1,373	5,243					738	2,979					635	2,264				
Number of beneficiaries	1,549,585	5,347,499					798,817	3,129,830					753,337	2,233,041				

Source: Mathematica's analysis of Medicare claims data from January 2016 through December 2021.

Notes: Although this table indicates which estimates are statistically significant, when we interpret evidence, we combine evidence from the magnitude of the effect, the p-values, findings on related outcomes, subgroups, sensitivity tests, and other data sources about model implementation.

[a] We report the actual, unadjusted averages in the baseline period, which are similar for the CPC+ and comparison groups because of matching. In the intervention periods, we computed the comparison group mean by subtracting the regression adjusted difference between the CPC+ and comparison means in each time period from the CPC+ mean in that same time period.

Table 5.J.2. ***(continued)***

[b] We regression-adjusted each impact estimate using a difference-in-differences analysis that reflects the difference of the average outcome for Medicare FFS beneficiaries attributed to CPC+ practices in the five years of CPC+ and the average outcome in the baseline year, relative to the same difference over time for Medicare FFS beneficiaries attributed to comparison practices, while controlling for beneficiary characteristics and practice fixed effects.

[c] We calculated percentage impacts relative to what the CPC+ mean would have been in PY 1 through PY 5 (separately and combined) in the absence of the intervention—that is, the unadjusted CPC+ mean minus the impact estimate.

*/**/*** Significantly different from zero at the 0.10/0.05/0.01 level, two-tailed test.

NA = not applicable, because the difference-in-differences impact estimate cannot be calculated at baseline.

CPC+ = Comprehensive Primary Care Plus; FFS = fee-for-service; PY = Program Year; SE = standard error.

Table 5.J.3. Regression-adjusted means and estimated impact of CPC+ on acute surgical and medical hospitalizations for attributed Medicare FFS beneficiaries over the five program years, Track 2

	Track 2—Overall						Track 2—SSP						Track 2—Non-SSP					
	CPC+ mean[a]	C mean[a]	Impact estimate[b] (SE)	Percentage impact[c]	90 percent confidence interval	*p*-Value	CPC+ mean[a]	C mean[a]	Impact estimate[b] (SE)	Percentage impact[c]	90 percent confidence interval	*p*-Value	CPC+ mean[a]	C mean[a]	Impact estimate[b] (SE)	Percentage impact[c]	90 percent confidence interval	*p*-Value
Monthly Medicare expenditures by service category (per beneficiary per month)																		
Total expenditures for acute hospital admissions																		
Baseline	$278	$281	NA	NA	NA	NA	$286	$285	NA	NA	NA	NA	$271	$278	NA	NA	NA	NA
PY 1	$284	$285	$2.7 ($2.1)	1.0%	(-$0.6, $6.1)	0.18	$293	$292	-$0.6 ($3.1)	-0.2%	(-$5.7, $4.5)	0.85	$278	$279	$5.4** ($2.7)	2.0%	($0.9, $9.9)	0.05
PY 2	$292	$294	$0.7 ($2.2)	0.3%	(-$3.0, $4.4)	0.75	$298	$300	-$3.5 ($3.5)	-1.2%	(-$9.3, $2.3)	0.32	$287	$289	$4.1 ($2.9)	1.4%	(-$0.7, $8.8)	0.16
PY 3	$298	$308	-$6.8*** ($2.3)	-2.2%	(-$10.7, -$3.0)	0.00	$306	$314	-$9.0** ($3.6)	-2.9%	(-$15.0, -$3.1)	0.01	$292	$303	-$5.2* ($3.1)	-1.7%	(-$10.2, -$0.2)	0.09
PY 4	$284	$292	-$5.3** ($2.6)	-1.8%	(-$9.6, -$0.9)	0.04	$291	$298	-$8.8** ($4.1)	-2.9%	(-$15.5, -$2.1)	0.03	$279	$286	-$0.9 ($3.2)	-0.3%	(-$6.1, $4.3)	0.77
PY 5	$298	$302	-$1.8 ($2.8)	-0.6%	(-$6.5, $2.8)	0.51	$305	$312	-$8.0* ($4.6)	-2.6%	(-$15.5, -$0.5)	0.08	$291	$295	$3.0 ($3.4)	1.0%	(-$2.5, $8.6)	0.37
PY 1 through 5	$291	$297	-$2.2 ($2.0)	-0.7%	(-$5.5, $1.1)	0.28	$299	$304	-$6.0* ($3.2)	-2.0%	(-$11.2, -$0.8)	0.06	$285	$291	$1.2 ($2.5)	0.4%	(-$2.9, $5.3)	0.64
Expenditures for acute surgical hospital admissions																		
Baseline	$147	$148	NA	NA	NA	NA	$151	$150	NA	NA	NA	NA	$143	$146	NA	NA	NA	NA
PY 1	$149	$146	$4.1*** ($1.5)	2.8%	($1.6, $6.6)	0.01	$152	$151	$0.5 ($2.3)	0.3%	(-$3.3, $4.3)	0.82	$147	$143	$7.0*** ($2.0)	5.0%	($3.7, $10.2)	0.00
PY 2	$151	$149	$3.4** ($1.6)	2.3%	($0.7, $6.0)	0.04	$152	$154	-$1.9 ($2.5)	-1.2%	(-$6.1, $2.3)	0.45	$150	$145	$7.5*** ($2.0)	5.3%	($4.2, $10.9)	0.00
PY 3	$155	$157	-$0.5 ($1.5)	-0.3%	(-$3.1, $2.0)	0.73	$158	$161	-$3.4 ($2.3)	-2.1%	(-$7.1, $0.3)	0.14	$152	$154	$1.7 ($2.1)	1.1%	(-$1.8, $5.1)	0.43
PY 4	$140	$141	$0.8 ($1.7)	0.6%	(-$2.0, $3.6)	0.65	$142	$145	-$3.1 ($2.6)	-2.1%	(-$7.4, $1.1)	0.23	$138	$136	$4.9** ($2.2)	3.7%	($1.3, $8.5)	0.02
PY 5	$143	$144	$0.7 ($1.7)	0.5%	(-$2.2, $3.6)	0.69	$145	$147	-$3.4 ($2.7)	-2.3%	(-$7.8, $1.1)	0.21	$142	$140	$4.7** ($2.3)	3.4%	($0.9, $8.5)	0.04
PY 1 through 5	$147	$147	$1.6 ($1.3)	1.1%	(-$0.5, $3.8)	0.20	$150	$151	-$2.3 ($2.0)	-1.5%	(-$5.5, $1.0)	0.25	$146	$144	$5.1*** ($1.7)	3.6%	($2.3, $7.9)	0.00

Table 5.J.3. ***(continued)***

	Track 2—Overall						Track 2—SSP						Track 2—Non-SSP					
	CPC+ mean[a]	C mean[a]	Impact estimate[b] (SE)	Percentage impact[c]	90 percent confidence interval	*p*-Value	CPC+ mean[a]	C mean[a]	Impact estimate[b] (SE)	Percentage impact[c]	90 percent confidence interval	*p*-Value	CPC+ mean[a]	C mean[a]	Impact estimate[b] (SE)	Percentage impact[c]	90 percent confidence interval	*p*-Value
Expenditures for acute medical hospital admissions																		
Baseline	$131	$133	NA	NA	NA	NA	$135	$134	NA	NA	NA	NA	$128	$131	NA	NA	NA	NA
PY 1	$135	$138	-$1.4 ($1.1)	-1.0%	(-$3.2, $0.5)	0.22	$141	$141	-$1.1 ($1.7)	-0.8%	(-$4.0, $1.7)	0.51	$131	$136	-$1.5 ($1.5)	-1.2%	(-$4.0, $0.9)	0.30
PY 2	$141	$145	-$2.6** ($1.2)	-1.8%	(-$4.7, -$0.6)	0.03	$146	$147	-$1.6 ($2.0)	-1.1%	(-$4.8, $1.6)	0.41	$137	$143	-$3.4** ($1.6)	-2.5%	(-$6.1, -$0.8)	0.03
PY 3	$143	$151	-$6.3*** ($1.4)	-4.2%	(-$8.6, -$4.0)	0.00	$148	$153	-$5.6** ($2.2)	-3.7%	(-$9.3, -$1.9)	0.01	$140	$150	-$6.8*** ($1.7)	-4.7%	(-$9.7, -$4.0)	0.00
PY 4	$144	$151	-$6.0*** ($1.5)	-4.0%	(-$8.6, -$3.5)	0.00	$148	$153	-$5.7** ($2.5)	-3.7%	(-$9.8, -$1.6)	0.02	$140	$149	-$5.8*** ($1.8)	-4.0%	(-$8.8, -$2.8)	0.00
PY 5	$154	$159	-$2.5 ($1.7)	-1.6%	(-$5.4, $0.3)	0.14	$161	$165	-$4.6 ($2.9)	-2.8%	(-$9.4, $0.1)	0.11	$150	$154	-$1.7 ($1.9)	-1.1%	(-$4.8, $1.5)	0.39
PY 1 through 5	$144	$149	-$3.8*** ($1.2)	-2.6%	(-$5.8, -$1.9)	0.00	$149	$152	-$3.7* ($1.9)	-2.5%	(-$6.9, -$0.6)	0.05	$140	$147	-$3.9*** ($1.4)	-2.7%	(-$6.2, -$1.6)	0.00
Annualized service use (per 1,000 beneficiaries per year)																		
Total annualized acute hospital admissions																		
Baseline	292	288	NA	NA	NA	NA	300	291	NA	NA	NA	NA	287	286	NA	NA	NA	NA
PY 1	292	289	-0.5 (1.6)	-0.2%	(-3.2, 2.1)	0.74	302	293	-0.4 (2.4)	-0.1%	(-4.3, 3.4)	0.85	285	285	-0.6 (2.2)	-0.2%	(-4.1, 3.0)	0.80
PY 2	289	286	-1.5 (1.7)	-0.5%	(-4.3, 1.3)	0.38	297	289	-0.1 (2.6)	0.0%	(-4.2, 4.1)	0.98	282	284	-2.7 (2.3)	-0.9%	(-6.5, 1.1)	0.25
PY 3	286	287	-4.8*** (1.8)	-1.7%	(-7.8, -1.8)	0.01	296	290	-2.1 (2.7)	-0.7%	(-6.6, 2.5)	0.45	278	285	-7.0*** (2.4)	-2.5%	(-11.0, -3.0)	0.00
PY 4	245	246	-4.8** (1.9)	-1.9%	(-8.0, -1.7)	0.01	253	248	-3.9 (3.0)	-1.5%	(-8.9, 1.1)	0.19	239	243	-5.1** (2.3)	-2.1%	(-8.9, -1.2)	0.03
PY 5	246	243	-1.9 (2.0)	-0.8%	(-5.1, 1.3)	0.34	256	250	-2.3 (3.1)	-0.9%	(-7.5, 2.9)	0.46	237	238	-1.9 (2.4)	-0.8%	(-5.9, 2.1)	0.44
PY 1 through 5	270	269	-2.7* (1.6)	-1.0%	(-5.3, -0.2)	0.08	280	273	-1.7 (2.4)	-0.6%	(-5.7, 2.2)	0.47	263	266	-3.5* (2.0)	-1.3%	(-6.8, -0.3)	0.08

Table 5.J.3. ***(continued)***

	Track 2—Overall						Track 2—SSP						Track 2—Non-SSP					
	CPC+ mean[a]	C mean[a]	Impact estimate[b] (SE)	Percentage impact[c]	90 percent confidence interval	*p*-Value	CPC+ mean[a]	C mean[a]	Impact estimate[b] (SE)	Percentage impact[c]	90 percent confidence interval	*p*-Value	CPC+ mean[a]	C mean[a]	Impact estimate[b] (SE)	Percentage impact[c]	90 percent confidence interval	*p*-Value
Annualized acute surgical hospital admissions																		
Baseline	90	88	NA	NA	NA	NA	91	88	NA	NA	NA	NA	89	87	NA	NA	NA	NA
PY 1	88	85	0.8 (0.7)	0.9%	(-0.4, 1.9)	0.27	89	87	-0.5 (1.0)	-0.6%	(-2.2, 1.1)	0.60	88	84	1.8* (0.9)	2.0%	(0.2, 3.3)	0.06
PY 2	87	84	0.8 (0.7)	1.0%	(-0.3, 1.9)	0.21	87	85	-0.3 (1.0)	-0.4%	(-2.0, 1.3)	0.75	87	83	1.7* (0.9)	2.1%	(0.3, 3.2)	0.05
PY 3	87	85	0.0 (0.7)	0.0%	(-1.1, 1.1)	0.95	88	86	0.0 (0.9)	0.0%	(-1.5, 1.6)	0.99	86	84	-0.1 (0.9)	-0.1%	(-1.7, 1.4)	0.90
PY 4	71	70	-0.9 (0.7)	-1.2%	(-2.0, 0.2)	0.19	71	71	-1.8* (1.0)	-2.4%	(-3.4, -0.1)	0.08	71	69	-0.2 (0.9)	-0.3%	(-1.8, 1.3)	0.79
PY 5	68	67	-0.6 (0.7)	-0.8%	(-1.7, 0.6)	0.42	69	68	-1.6 (1.0)	-2.2%	(-3.3, 0.2)	0.14	68	65	0.4 (1.0)	0.6%	(-1.1, 2.0)	0.66
PY 1 through 5	80	78	0.0 (0.6)	0.0%	(-0.9, 1.0)	0.96	80	79	-0.8 (0.8)	-1.0%	(-2.1, 0.5)	0.32	79	77	0.7 (0.8)	0.9%	(-0.6, 2.0)	0.36
Annualized acute medical hospital admissions																		
Baseline	202	201	NA	NA	NA	NA	209	203	NA	NA	NA	NA	198	199	NA	NA	NA	NA
PY 1	204	203	-1.3 (1.3)	-0.6%	(-3.5, 0.9)	0.34	213	207	0.1 (2.0)	0.0%	(-3.3, 3.4)	0.97	197	200	-2.3 (1.8)	-1.2%	(-5.3, 0.6)	0.19
PY 2	202	203	-2.4 (1.4)	-1.2%	(-4.7, 0.0)	0.10	210	204	0.3 (2.2)	0.1%	(-3.3, 3.8)	0.90	196	202	-4.4** (1.9)	-2.2%	(-7.6, -1.3)	0.02
PY 3	199	202	-4.8*** (1.5)	-2.3%	(-7.3, -2.3)	0.00	208	204	-2.1 (2.4)	-1.0%	(-6.0, 1.8)	0.38	192	200	-6.9*** (2.0)	-3.5%	(-10.2, -3.7)	0.00
PY 4	174	176	-3.9** (1.6)	-2.2%	(-6.6, -1.3)	0.01	182	178	-2.2 (2.6)	-1.2%	(-6.5, 2.1)	0.41	168	174	-4.8** (1.9)	-2.8%	(-8.0, -1.7)	0.01
PY 5	178	177	-1.3 (1.6)	-0.7%	(-4.0, 1.4)	0.43	188	182	-0.8 (2.7)	-0.4%	(-5.1, 3.6)	0.78	169	173	-2.3 (2.0)	-1.4%	(-5.7, 1.0)	0.25
PY 1 through 5	191	191	-2.8** (1.3)	-1.4%	(-4.9, -0.6)	0.03	199	194	-0.9 (2.1)	-0.5%	(-4.3, 2.4)	0.65	184	189	-4.3*** (1.6)	-2.3%	(-6.9, -1.6)	0.01
Unweighted sample sizes																		
Number of practices	1,515	3,783					636	1,817					879	1,966				
Number of beneficiaries	1,896,880	4,507,499					847,208	2,257,322					1,053,634	2,261,852				

Source: Mathematica's analysis of Medicare claims data from January 2016 through December 2021.

Notes: Although this table indicates which estimates are statistically significant, when we interpret evidence, we combine evidence from the magnitude of the effect, the p-values, findings on related outcomes, subgroups, sensitivity tests, and other data sources about model implementation.

[a] We report the actual, unadjusted averages in the baseline period which are similar for the CPC+ and comparison groups because of matching. In the intervention periods, we computed the comparison group mean by subtracting the regression adjusted difference between the CPC+ and comparison means in each time period from the CPC+ mean in that same time period.

Table 5.J.3. ***(continued)***

[b] We regression-adjusted each impact estimate using a difference-in-differences analysis that reflects the difference of the average outcome for Medicare FFS beneficiaries attributed to CPC+ practices in the five years of CPC+ and the average outcome in the baseline year, relative to the same difference over time for Medicare FFS beneficiaries attributed to comparison practices, while controlling for beneficiary characteristics and practice fixed effects.

[c] We calculated percentage impacts relative to what the CPC+ mean would have been in PY 1 through PY 5 (separately and combined) in the absence of the intervention—that is, the unadjusted CPC+ mean minus the impact estimate.

*/**/*** Significantly different from zero at the 0.10/0.05/0.01 level, two-tailed test.

NA = not applicable, because the difference-in-differences impact estimate cannot be calculated at baseline.

CPC+ = Comprehensive Primary Care Plus; FFS = fee-for-service; PY = Program Year; SE = standard error.

C. Impact on severity of surgical and medical hospitalizations

In both tracks, admissions without an MCC or CC drove the reductions in the number of medical admissions. For both CPC+ and comparison practices, the number of medical admissions without an MCC or CC declined from the baseline year, but this decline was greater for CPC+ practices, resulting in annual average reductions of 1.6 and 1.5 hospitalizations per 1,000 beneficiaries in Tracks 1 and 2, respectively (2.6 percent, $p = 0.010$ for Track 1; and 2.4 percent, $p = 0.013$ for Track 2; Tables 5.J.4 and 5.J.5). The magnitude of these estimates indicates that reductions in acute medical admissions without an MCC or CC accounted for more than half of the reduction in all-cause acute hospitalizations from CPC+ (1.6 out of 2.5 reduced admissions for Track 1 and 1.5 out of 2.7 reduced admissions for Track 2), despite making up less than 30 percent of all hospital admissions. In Track 1, CPC+ reduced medical admissions without an MCC or CC every program year, with the largest effect in the third program year, at 3.0 percent ($p < 0.01$; Table 5.J.4). Similarly, for Track 2, there were reductions in medical hospital admissions without an MCC or CC each year, except PY 1, and the largest reduction was also in PY 3, at 3.2 percent ($p < 0.01$; Table 5.J.5). Although there were no annual average reductions in medical admissions with an MCC or with a CC, there were significant reductions in medical admissions with an MCC for PY 4 in both tracks.

Despite the lack of observed effects on higher severity medical admissions, the reductions in expenditures for acute medical admissions came from reduced expenditures for all severity categories, with reduced expenditures for medical admissions with an MCC contributing the most. Across the five program years in each track, there were reductions in expenditures for medical admissions without an MCC or CC (2.2 and 2.1 percent reductions for Track 1 and 2, respectively), with an CC (2.2 and 2.3 percent for Track 1 and 2 respectively), and with an MCC (2.2 and 2.9 percent for Track 1 and 2, respectively). For both tracks, the magnitude of the average annual reduction was largest for medical admissions with an MCC compared with the two other severity levels (Tables 5.J.4 and 5.J.5).

For surgical hospital admissions, there is little evidence CPC+ affected the number of hospitalizations in any of the severity categories or the expenditures for any category of surgical admissions. Over the course of the model, no category showed significant average effects for the rate of hospitalizations or hospital expenditures, and there were only a handful of significant estimates in individual program years (favorable and unfavorable; Tables 5.J.6 and 5.J.7).

In both tracks, non-SSP practices reduced the least severe hospitalizations, and there is some evidence SSP practices were able to reduce more complex admissions. Over the five program years, non-SSP practices reduced acute medical admissions without an MCC or CC by 3.5 percent in Track 1 ($p = 0.02$) and 4.0 percent in Track 2 ($p < 0.01$). In Track 2, non-SSP practices also reduced medical admissions with a CC multiple years, with an average annual reduction of 3.9 percent ($p < 0.01$). For Track 1 SSP practices, there were reductions in medical admissions with an MCC, with an average annual reduction of 2.2 percent ($p = 0.03$; Table 5.J.4). In both tracks, SSP reduced surgical admissions without an MCC or CC in the last two program years (Tables 5.J.6 and 5.J.7). For Track 2 SSP practices, there are reductions in surgical admissions with an MCC in each program year, with an average annual reduction of 5.5 percent ($p < 0.01$; Table 5.J.7).

Turning to hospital expenditures, Track 1 SSP practices had reductions along multiple categories of severity; for Track 2, both SSP and non-SSP practices had reductions in multiple severity levels. In Track 1, the reductions in lower-complexity medical admissions among non-SSP practices did not translate into statistically significant reduced average annual medical expenditures for medical admissions without an

MCC or CC (Table 5.J.4). Track 1 SSP practices reduced expenditures for medical admissions with an MCC (3.4 percent, $p < 0.01$) and with a CC (4.1 percent, $p < 0.01$). Track 1 SSP practices also reduced expenditures for surgical admissions with an MCC (2.9 percent, $p = 0.096$). However, this did not translate into reductions in annual average expenditures for overall surgical admissions (Table 5.J.6). For Track 2, CPC+ reduced medical expenditures for both SSP and non-SSP practices over the five years of the model, reducing expenditures for all severity categories for non-SSP practices and reducing only medical admissions with an MCC for SSP practices (Table 5.J.5). In addition, for Track 2 SSP practices, CPC+ reduced expenditures for surgical admissions with an MCC (4.7 percent, $p = 0.02$; Table 5.J.7).

Table 5.J.4. Regression-adjusted means and estimated impact of CPC+ on types of medical hospitalizations for attributed Medicare FFS beneficiaries over the five program years, Track 1

	Track 1—Overall						Track 1—SSP						Track 1—Non-SSP					
	CPC+ mean[a]	C mean[a]	Impact estimate[b] (SE)	Percentage impact[c]	90 percent confidence interval	*p*-Value	CPC+ mean[a]	C mean[a]	Impact estimate[b] (SE)	Percentage impact[c]	90 percent confidence interval	*p*-Value	CPC+ mean[a]	C mean[a]	Impact estimate[b] (SE)	Percentage impact[c]	90 percent confidence interval	*p*-Value
Monthly Medicare expenditures by service category (per beneficiary per month)																		
Expenditures for acute medical hospital admissions with an MCC																		
Baseline	$65	$68	NA	NA	NA	NA	$66	$68	NA	NA	NA	NA	$63	$67	NA	NA	NA	NA
PY 1	$71	$75	-$0.7 ($0.7)	-1.0%	(-$2.0, $0.5)	0.33	$73	$76	-$0.8 ($1.0)	-1.1%	(-$2.4, $0.7)	0.38	$70	$74	-$0.6 ($1.2)	-0.8%	(-$2.5, $1.3)	0.62
PY 2	$77	$80	-$0.2 ($0.8)	-0.3%	(-$1.5, $1.1)	0.80	$79	$81	-$0.2 ($1.1)	-0.2%	(-$1.9, $1.6)	0.86	$74	$78	-$0.2 ($1.2)	-0.3%	(-$2.2, $1.8)	0.87
PY 3	$80	$85	-$2.2** ($0.9)	-2.7%	(-$3.7, -$0.7)	0.01	$82	$88	-$3.6*** ($1.3)	-4.2%	(-$5.6, -$1.5)	0.00	$78	$82	-$0.6 ($1.3)	-0.8%	(-$2.7, $1.4)	0.61
PY 4	$85	$92	-$4.2*** ($1.0)	-4.7%	(-$5.8, -$2.6)	0.00	$89	$96	-$5.7*** ($1.4)	-6.0%	(-$7.9, -$3.4)	0.00	$82	$88	-$2.7* ($1.4)	-3.2%	(-$5.1, -$0.4)	0.06
PY 5	$94	$99	-$2.1* ($1.1)	-2.1%	(-$3.8, -$0.3)	0.05	$99	$106	-$5.2*** ($1.5)	-5.0%	(-$7.7, -$2.7)	0.00	$89	$93	$0.3 ($1.5)	0.3%	(-$2.2, $2.8)	0.85
PY 1 through 5	$82	$87	-$1.9*** ($0.7)	-2.2%	(-$3.0, -$0.7)	0.01	$85	$90	-$3.0*** ($1.0)	-3.4%	(-$4.6, -$1.4)	0.00	$79	$83	-$0.7 ($1.0)	-0.9%	(-$2.5, $1.0)	0.48
Expenditures for acute medical hospital admissions with a CC																		
Baseline	$27	$28	NA	NA	NA	NA	$28	$28	NA	NA	NA	NA	$26	$27	NA	NA	NA	NA
PY 1	$25	$26	-$0.7** ($0.3)	-2.8%	(-$1.3, -$0.1)	0.04	$26	$27	-$0.8 ($0.5)	-2.9%	(-$1.6, $0.0)	0.11	$24	$26	-$0.7 ($0.5)	-2.7%	(-$1.5, $0.2)	0.18
PY 2	$25	$27	-$0.6* ($0.4)	-2.4%	(-$1.2, $0.0)	0.09	$26	$27	-$0.8 ($0.5)	-3.0%	(-$1.6, $0.0)	0.10	$25	$27	-$0.4 ($0.5)	-1.7%	(-$1.3, $0.4)	0.42
PY 3	$26	$27	-$0.5 ($0.4)	-1.8%	(-$1.1, $0.2)	0.22	$26	$28	-$1.1** ($0.5)	-3.9%	(-$1.9, -$0.3)	0.03	$25	$27	$0.2 ($0.6)	0.7%	(-$0.8, $1.1)	0.77
PY 4	$24	$26	-$0.9** ($0.4)	-3.4%	(-$1.5, -$0.2)	0.04	$24	$26	-$1.8*** ($0.5)	-7.1%	(-$2.7, -$0.9)	0.00	$24	$25	$0.1 ($0.6)	0.3%	(-$1.0, $1.2)	0.90
PY 5	$25	$26	-$0.2 ($0.4)	-0.7%	(-$0.9, $0.5)	0.67	$25	$26	-$1.0* ($0.6)	-3.9%	(-$2.0, $0.0)	0.10	$24	$25	$0.7 ($0.6)	2.9%	(-$0.3, $1.7)	0.24
PY 1 through 5	$25	$26	-$0.6* ($0.3)	-2.2%	(-$1.1, -$0.1)	0.06	$25	$27	-$1.1*** ($0.4)	-4.1%	(-$1.8, -$0.4)	0.01	$24	$26	$0.0 ($0.5)	-0.1%	(-$0.8, $0.7)	0.97

Table 5.J.4. ***(continued)***

	Track 1—Overall						Track 1—SSP						Track 1—Non-SSP					
	CPC+ mean[a]	C mean[a]	Impact estimate[b] (SE)	Percentage impact[c]	90 percent confidence interval	*p*-Value	CPC+ mean[a]	C mean[a]	Impact estimate[b] (SE)	Percentage impact[c]	90 percent confidence interval	*p*-Value	CPC+ mean[a]	C mean[a]	Impact estimate[b] (SE)	Percentage impact[c]	90 percent confidence interval	*p*-Value
Expenditures for acute medical hospital admissions without an MCC or CC																		
Baseline	$36	$38	NA	NA	NA	NA	$36	$38	NA	NA	NA	NA	$35	$38	NA	NA	NA	NA
PY 1	$35	$37	$0.0 ($0.4)	0.1%	(-$0.6, $0.7)	0.93	$35	$37	-$0.2 ($0.6)	-0.5%	(-$1.1, $0.7)	0.75	$34	$36	$0.3 ($0.6)	0.8%	(-$0.7, $1.3)	0.63
PY 2	$34	$37	-$1.1** ($0.4)	-3.2%	(-$1.9, -$0.4)	0.01	$35	$38	-$1.3** ($0.6)	-3.5%	(-$2.3, -$0.3)	0.04	$33	$37	-$1.0 ($0.6)	-2.8%	(-$2.0, $0.1)	0.13
PY 3	$35	$38	-$1.2** ($0.5)	-3.2%	(-$1.9, -$0.4)	0.01	$36	$39	-$1.1* ($0.6)	-3.0%	(-$2.2, -$0.1)	0.08	$34	$38	-$1.2* ($0.7)	-3.4%	(-$2.3, -$0.1)	0.08
PY 4	$32	$35	-$0.8 ($0.6)	-2.4%	(-$1.7, $0.2)	0.17	$33	$36	-$0.7 ($0.7)	-2.2%	(-$1.9, $0.5)	0.32	$31	$35	-$1.0 ($0.9)	-3.0%	(-$2.5, $0.6)	0.30
PY 5	$33	$35	-$0.6 ($0.5)	-1.8%	(-$1.5, $0.3)	0.28	$35	$36	$0.5 ($0.8)	1.5%	(-$0.8, $1.8)	0.52	$31	$35	-$1.7** ($0.8)	-5.2%	(-$3.0, -$0.4)	0.03
PY 1 through 5	$34	$37	-$0.7** ($0.4)	-2.2%	(-$1.4, -$0.1)	0.05	$35	$37	-$0.6 ($0.5)	-1.7%	(-$1.4, $0.2)	0.22	$33	$36	-$0.9 ($0.6)	-2.8%	(-$1.9, $0.0)	0.10
Annualized service use (per 1,000 beneficiaries per year)																		
Annualized acute medical hospital admissions with an MCC																		
Baseline	75	75	NA	NA	NA	NA	75	75	NA	NA	NA	NA	74	75	NA	NA	NA	NA
PY 1	82	83	-0.2 (0.7)	-0.2%	(-1.4, 1.0)	0.81	82	83	-1.4 (0.9)	-1.7%	(-2.9, 0.1)	0.13	82	82	1.1 (1.1)	1.4%	(-0.7, 3.0)	0.31
PY 2	85	85	0.1 (0.8)	0.2%	(-1.2, 1.5)	0.86	85	85	-0.3 (1.0)	-0.4%	(-2.0, 1.4)	0.76	84	84	0.6 (1.2)	0.7%	(-1.4, 2.7)	0.62
PY 3	86	86	-0.6 (0.8)	-0.7%	(-2.0, 0.7)	0.44	86	88	-2.3** (1.1)	-2.6%	(-4.1, -0.5)	0.04	85	85	1.2 (1.3)	1.4%	(-0.9, 3.3)	0.35
PY 4	82	85	-2.3*** (0.9)	-2.7%	(-3.8, -0.9)	0.01	84	88	-4.1*** (1.2)	-4.6%	(-6.1, -2.1)	0.00	81	82	-0.3 (1.3)	-0.4%	(-2.5, 1.8)	0.80
PY 5	87	87	0.2 (0.9)	0.2%	(-1.3, 1.7)	0.83	90	91	-1.8 (1.2)	-2.0%	(-3.8, 0.2)	0.14	84	83	2.3* (1.4)	2.8%	(0.0, 4.5)	0.10
PY 1 through 5	84	85	-0.5 (0.7)	-0.6%	(-1.7, 0.6)	0.42	86	87	-1.9** (0.9)	-2.2%	(-3.4, -0.5)	0.03	83	83	1.0 (1.0)	1.2%	(-0.7, 2.7)	0.34

Table 5.J.4. ***(continued)***

	Track 1—Overall						Track 1—SSP						Track 1—Non-SSP					
	CPC+ mean[a]	C mean[a]	Impact estimate[b] (SE)	Percentage impact[c]	90 percent confidence interval	*p*-Value	CPC+ mean[a]	C mean[a]	Impact estimate[b] (SE)	Percentage impact[c]	90 percent confidence interval	*p*-Value	CPC+ mean[a]	C mean[a]	Impact estimate[b] (SE)	Percentage impact[c]	90 percent confidence interval	*p*-Value
Annualized acute medical hospital admissions with a CC																		
Baseline	48	49	NA	NA	NA	NA	49	49	NA	NA	NA	NA	47	49	NA	NA	NA	NA
PY 1	45	46	-0.4 (0.5)	-0.9%	(-1.3, 0.5)	0.46	46	46	-0.2 (0.7)	-0.5%	(-1.4, 0.9)	0.76	44	46	-0.6 (0.8)	-1.3%	(-1.9, 0.7)	0.46
PY 2	45	46	-0.3 (0.5)	-0.7%	(-1.2, 0.6)	0.56	45	46	-0.6 (0.7)	-1.2%	(-1.7, 0.6)	0.42	44	46	0.0 (0.8)	-0.1%	(-1.3, 1.3)	0.96
PY 3	44	45	0.2 (0.6)	0.4%	(-0.7, 1.1)	0.72	45	45	-0.6 (0.7)	-1.4%	(-1.8, 0.5)	0.36	44	44	1.1 (0.9)	2.6%	(-0.3, 2.6)	0.20
PY 4	36	37	-0.3 (0.6)	-0.9%	(-1.3, 0.6)	0.54	36	38	-1.6** (0.7)	-4.2%	(-2.8, -0.4)	0.03	37	37	0.9 (0.9)	2.4%	(-0.5, 2.3)	0.31
PY 5	36	37	-0.4 (0.5)	-1.2%	(-1.3, 0.5)	0.42	36	37	-1.2 (0.7)	-3.1%	(-2.4, 0.1)	0.12	35	36	0.5 (0.8)	1.5%	(-0.8, 1.8)	0.53
PY 1 through 5	41	42	-0.2 (0.5)	-0.6%	(-1.0, 0.5)	0.59	41	42	-0.8 (0.6)	-1.9%	(-1.8, 0.2)	0.17	41	42	0.4 (0.7)	1.0%	(-0.8, 1.5)	0.58
Annualized acute medical hospital admissions without an MCC or CC																		
Baseline	77	77	NA	NA	NA	NA	76	77	NA	NA	NA	NA	78	78	NA	NA	NA	NA
PY 1	73	74	-1.1* (0.7)	-1.5%	(-2.2, 0.0)	0.10	72	75	-1.1 (0.9)	-1.5%	(-2.6, 0.3)	0.20	73	74	-1.0 (1.0)	-1.3%	(-2.6, 0.6)	0.31
PY 2	69	71	-1.8** (0.7)	-2.6%	(-3.0, -0.6)	0.01	69	71	-1.5 (1.0)	-2.1%	(-3.1, 0.1)	0.13	69	70	-2.1* (1.1)	-3.0%	(-3.9, -0.3)	0.05
PY 3	67	70	-2.1*** (0.8)	-3.0%	(-3.4, -0.8)	0.01	68	71	-1.4 (1.0)	-2.1%	(-3.0, 0.2)	0.14	67	69	-2.8** (1.2)	-4.0%	(-4.7, -0.8)	0.02
PY 4	53	55	-1.4* (0.8)	-2.5%	(-2.7, 0.0)	0.10	53	55	-0.4 (1.0)	-0.7%	(-2.0, 1.3)	0.72	52	54	-2.2* (1.3)	-4.0%	(-4.3, -0.1)	0.09
PY 5	53	55	-1.7** (0.8)	-3.2%	(-3.0, -0.4)	0.03	54	55	0.2 (1.0)	0.3%	(-1.5, 1.9)	0.87	51	54	-3.0** (1.2)	-5.5%	(-4.9, -1.0)	0.01
PY 1 through 5	62	64	-1.6*** (0.6)	-2.6%	(-2.7, -0.6)	0.01	63	65	-0.9 (0.8)	-1.4%	(-2.2, 0.4)	0.27	62	64	-2.2** (1.0)	-3.5%	(-3.8, -0.6)	0.02
Unweighted sample sizes																		
Number of practices	1,373	5,243					738	2,979					635	2,264				
Number of beneficiaries	1,549,585	5,347,499					798,817	3,129,830					753,337	2,233,041				

Source: Mathematica's analysis of Medicare claims data from January 2016 through December 2021.

Notes: Although this table indicates which estimates are statistically significant, when we interpret evidence, we combine evidence from the magnitude of the effect, the p-values, findings on related outcomes, subgroups, sensitivity tests, and other data sources about model implementation.

[a] We report the actual, unadjusted averages in the baseline period which are similar for the CPC+ and comparison groups because of matching. In the intervention periods, we computed the comparison group mean by subtracting the regression adjusted difference between the CPC+ and comparison means in each time period from the CPC+ mean in that same time period.

Table 5.J.4. ***(continued)***

[b] We regression-adjusted each impact estimate using a difference-in-differences analysis that reflects the difference of the average outcome for Medicare FFS beneficiaries attributed to CPC+ practices in the five years of CPC+ and the average outcome in the baseline year, relative to the same difference over time for Medicare FFS beneficiaries attributed to comparison practices, while controlling for beneficiary characteristics and practice fixed effects.

[c] We calculated percentage impacts relative to what the CPC+ mean would have been in PY 1 through PY 5 (separately and combined) in the absence of the intervention—that is, the unadjusted CPC+ mean minus the impact estimate.

*/**/*** Significantly different from zero at the 0.10/0.05/0.01 level, two-tailed test.

NA = not applicable, because the difference-in-differences impact estimate cannot be calculated at baseline.

CC = complication or comorbidity; CPC+ = Comprehensive Primary Care Plus; FFS = fee-for-service; MCC = major complication or comorbidity; PY = Program Year; SE = standard error.

Table 5.J.5. Regression-adjusted means and estimated impact of CPC+ on types of medical hospitalizations for attributed Medicare FFS beneficiaries over the five program years, Track 2

	Track 2—Overall						Track 2—SSP						Track 2—Non-SSP					
	CPC+ mean[a]	C mean[a]	Impact estimate[b] (SE)	Percentage impact[c]	90 percent confidence interval	*p*-Value	CPC+ mean[a]	C mean[a]	Impact estimate[b] (SE)	Percentage impact[c]	90 percent confidence interval	*p*-Value	CPC+ mean[a]	C mean[a]	Impact estimate[b] (SE)	Percentage impact[c]	90 percent confidence interval	*p*-Value
Monthly Medicare expenditures by service category (per beneficiary per month)																		
Expenditures for acute medical hospital admissions with an MCC																		
Baseline	$67	$67	NA	NA	NA	NA	$68	$68	NA	NA	NA	NA	$65	$66	NA	NA	NA	NA
PY 1	$74	$75	-$0.6 ($0.8)	-0.8%	(-$1.9, $0.7)	0.46	$76	$77	-$0.7 ($1.2)	-1.0%	(-$2.7, $1.2)	0.53	$72	$73	-$0.5 ($1.1)	-0.6%	(-$2.3, $1.3)	0.67
PY 2	$79	$80	-$1.0 ($0.9)	-1.3%	(-$2.5, $0.5)	0.26	$81	$82	-$1.4 ($1.4)	-1.7%	(-$3.7, $0.9)	0.31	$77	$79	-$0.7 ($1.2)	-0.9%	(-$2.7, $1.2)	0.55
PY 3	$81	$85	-$3.5*** ($1.0)	-4.1%	(-$5.1, -$1.8)	0.00	$83	$87	-$4.2*** ($1.6)	-4.9%	(-$6.8, -$1.6)	0.01	$80	$83	-$2.9** ($1.3)	-3.5%	(-$5.0, -$0.8)	0.02
PY 4	$86	$91	-$5.0*** ($1.1)	-5.5%	(-$6.8, -$3.1)	0.00	$88	$93	-$5.5*** ($1.8)	-5.9%	(-$8.4, -$2.6)	0.00	$84	$89	-$4.3*** ($1.3)	-4.8%	(-$6.5, -$2.0)	0.00
PY 5	$95	$98	-$2.2* ($1.2)	-2.3%	(-$4.3, -$0.2)	0.07	$98	$103	-$4.9** ($2.1)	-4.8%	(-$8.4, -$1.4)	0.02	$93	$94	-$0.5 ($1.4)	-0.6%	(-$2.8, $1.7)	0.69
PY 1 through 5	$83	$86	-$2.5*** ($0.8)	-2.9%	(-$3.8, -$1.1)	0.00	$86	$89	-$3.3** ($1.3)	-3.8%	(-$5.6, -$1.1)	0.01	$82	$84	-$1.8* ($1.0)	-2.1%	(-$3.4, -$0.1)	0.08
Expenditures for acute medical hospital admissions with a CC																		
Baseline	$27	$28	NA	NA	NA	NA	$28	$28	NA	NA	NA	NA	$27	$27	NA	NA	NA	NA
PY 1	$26	$26	-$0.3 ($0.4)	-1.1%	(-$0.9, $0.3)	0.43	$27	$27	$0.3 ($0.5)	1.1%	(-$0.5, $1.1)	0.57	$25	$26	-$0.7 ($0.5)	-2.8%	(-$1.5, $0.1)	0.14
PY 2	$26	$27	-$0.9** ($0.4)	-3.3%	(-$1.5, -$0.3)	0.02	$27	$27	$0.1 ($0.5)	0.5%	(-$0.8, $1.0)	0.81	$25	$27	-$1.7*** ($0.5)	-6.4%	(-$2.5, -$0.8)	0.00
PY 3	$26	$28	-$1.1*** ($0.4)	-4.2%	(-$1.8, -$0.5)	0.01	$27	$27	$0.0 ($0.6)	-0.1%	(-$1.0, $1.0)	0.98	$25	$28	-$2.0*** ($0.6)	-7.5%	(-$3.0, -$1.1)	0.00
PY 4	$25	$26	-$0.6 ($0.4)	-2.5%	(-$1.4, $0.1)	0.15	$26	$26	-$0.4 ($0.7)	-1.5%	(-$1.6, $0.8)	0.57	$24	$25	-$0.8 ($0.6)	-3.1%	(-$1.7, $0.2)	0.17
PY 5	$25	$26	$0.0 ($0.5)	-0.1%	(-$0.8, $0.7)	0.94	$27	$26	$0.7 ($0.7)	2.9%	(-$0.4, $1.9)	0.29	$25	$26	-$0.6 ($0.6)	-2.4%	(-$1.6, $0.4)	0.33
PY 1 through 5	$26	$27	-$0.6* ($0.3)	-2.3%	(-$1.1, -$0.1)	0.06	$27	$27	$0.1 ($0.5)	0.5%	(-$0.7, $0.9)	0.76	$25	$26	-$1.2*** ($0.4)	-4.7%	(-$1.9, -$0.5)	0.01

Table 5.J.5. ***(continued)***

	Track 2—Overall						Track 2—SSP						Track 2—Non-SSP					
	CPC+ mean[a]	C mean[a]	Impact estimate[b] (SE)	Percentage impact[c]	90 percent confidence interval	*p*-Value	CPC+ mean[a]	C mean[a]	Impact estimate[b] (SE)	Percentage impact[c]	90 percent confidence interval	*p*-Value	CPC+ mean[a]	C mean[a]	Impact estimate[b] (SE)	Percentage impact[c]	90 percent confidence interval	*p*-Value
Expenditures for acute medical hospital admissions without an MCC or CC																		
Baseline	$37	$38	NA	NA	NA	NA	$38	$38	NA	NA	NA	NA	$36	$38	NA	NA	NA	NA
PY 1	$36	$37	-$0.5 ($0.5)	-1.4%	(-$1.3, $0.3)	0.28	$38	$38	-$0.7 ($0.8)	-1.8%	(-$1.9, $0.6)	0.37	$35	$37	-$0.4 ($0.6)	-1.0%	(-$1.4, $0.6)	0.56
PY 2	$36	$38	-$0.8 ($0.5)	-2.1%	(-$1.6, $0.1)	0.13	$38	$37	-$0.3 ($0.7)	-0.9%	(-$1.6, $0.9)	0.65	$35	$38	-$1.1 ($0.7)	-2.9%	(-$2.1, $0.0)	0.11
PY 3	$36	$39	-$1.7*** ($0.5)	-4.5%	(-$2.5, -$0.8)	0.00	$38	$39	-$1.4* ($0.8)	-3.5%	(-$2.6, -$0.1)	0.08	$35	$39	-$1.9*** ($0.7)	-5.2%	(-$3.0, -$0.8)	0.01
PY 4	$33	$34	-$0.4 ($0.6)	-1.3%	(-$1.3, $0.5)	0.45	$34	$34	$0.2 ($0.9)	0.6%	(-$1.3, $1.7)	0.82	$32	$35	-$0.8 ($0.7)	-2.4%	(-$2.0, $0.4)	0.27
PY 5	$34	$35	-$0.3 ($0.6)	-0.8%	(-$1.3, $0.7)	0.67	$36	$36	-$0.5 ($1.0)	-1.3%	(-$2.1, $1.2)	0.63	$32	$35	-$0.5 ($0.8)	-1.6%	(-$1.8, $0.7)	0.49
PY 1 through 5	$35	$37	-$0.8* ($0.4)	-2.1%	(-$1.5, -$0.1)	0.08	$37	$37	-$0.6 ($0.7)	-1.5%	(-$1.7, $0.6)	0.41	$34	$37	-$1.0* ($0.5)	-2.8%	(-$1.9, -$0.1)	0.08
Annualized service use (per 1,000 beneficiaries per year)																		
Annualized acute medical hospital admissions with an MCC																		
Baseline	75	74	NA	NA	NA	NA	76	75	NA	NA	NA	NA	74	74	NA	NA	NA	NA
PY 1	83	82	0.6 (0.8)	0.8%	(-0.6, 1.9)	0.41	85	84	0.5 (1.1)	0.6%	(-1.3, 2.4)	0.65	82	81	0.7 (1.0)	0.9%	(-1.0, 2.4)	0.50
PY 2	86	85	-0.2 (0.8)	-0.3%	(-1.6, 1.1)	0.79	87	87	-0.5 (1.2)	-0.6%	(-2.5, 1.5)	0.67	85	84	0.0 (1.1)	0.0%	(-1.8, 1.8)	1.00
PY 3	86	87	-1.4 (0.9)	-1.6%	(-2.9, 0.1)	0.11	88	88	-1.7 (1.4)	-1.9%	(-4.0, 0.7)	0.24	84	85	-1.2 (1.1)	-1.4%	(-3.1, 0.6)	0.28
PY 4	83	84	-1.9** (1.0)	-2.3%	(-3.6, -0.3)	0.05	85	86	-2.1 (1.5)	-2.5%	(-4.6, 0.3)	0.16	81	82	-1.4 (1.2)	-1.7%	(-3.4, 0.6)	0.24
PY 5	87	86	0.7 (1.0)	0.8%	(-1.0, 2.4)	0.50	90	90	-0.8 (1.7)	-0.9%	(-3.5, 1.9)	0.62	85	83	1.4 (1.2)	1.7%	(-0.6, 3.5)	0.25
PY 1 through 5	85	85	-0.5 (0.8)	-0.5%	(-1.7, 0.8)	0.53	87	87	-0.9 (1.2)	-1.1%	(-2.9, 1.0)	0.42	83	83	-0.1 (1.0)	-0.2%	(-1.7, 1.4)	0.89

Table 5.J.5. ***(continued)***

	Track 2—Overall						Track 2—SSP						Track 2—Non-SSP					
	CPC+ mean[a]	C mean[a]	Impact estimate[b] (SE)	Percentage impact[c]	90 percent confidence interval	*p*-Value	CPC+ mean[a]	C mean[a]	Impact estimate[b] (SE)	Percentage impact[c]	90 percent confidence interval	*p*-Value	CPC+ mean[a]	C mean[a]	Impact estimate[b] (SE)	Percentage impact[c]	90 percent confidence interval	*p*-Value
Annualized acute medical hospital admissions with a CC																		
Baseline	49	49	NA	NA	NA	NA	51	50	NA	NA	NA	NA	48	48	NA	NA	NA	NA
PY 1	46	46	-0.8 (0.5)	-1.7%	(-1.7, 0.1)	0.13	48	47	0.3 (0.8)	0.6%	(-1.0, 1.6)	0.71	44	46	-1.6** (0.7)	-3.6%	(-2.8, -0.5)	0.02
PY 2	45	46	-0.7 (0.6)	-1.6%	(-1.6, 0.2)	0.19	48	46	0.6 (0.8)	1.4%	(-0.7, 2.0)	0.44	44	46	-1.8** (0.7)	-4.0%	(-3.0, -0.6)	0.01
PY 3	44	45	-1.1* (0.6)	-2.3%	(-2.0, -0.1)	0.06	47	45	0.4 (0.8)	0.8%	(-1.0, 1.8)	0.65	42	45	-2.2*** (0.7)	-4.9%	(-3.4, -1.0)	0.00
PY 4	37	37	-0.5 (0.6)	-1.4%	(-1.5, 0.4)	0.34	39	37	0.2 (0.9)	0.4%	(-1.3, 1.7)	0.85	36	37	-1.0 (0.7)	-2.7%	(-2.1, 0.2)	0.15
PY 5	36	37	-0.6 (0.6)	-1.7%	(-1.6, 0.3)	0.27	39	37	0.4 (0.9)	1.2%	(-1.0, 1.9)	0.62	34	36	-1.4* (0.8)	-3.8%	(-2.6, -0.1)	0.07
PY 1 through 5	42	42	-0.8 (0.5)	-1.8%	(-1.5, 0.0)	0.10	44	42	0.4 (0.7)	0.9%	(-0.8, 1.5)	0.58	40	42	-1.6*** (0.6)	-3.9%	(-2.6, -0.6)	0.01
Annualized acute medical hospital admissions without an MCC or CC																		
Baseline	78	78	NA	NA	NA	NA	82	78	NA	NA	NA	NA	76	77	NA	NA	NA	NA
PY 1	75	75	-1.1 (0.7)	-1.5%	(-2.2, 0.0)	0.11	79	76	-0.7 (1.0)	-0.9%	(-2.4, 1.0)	0.49	71	74	-1.4 (0.9)	-1.9%	(-2.9, 0.1)	0.13
PY 2	71	72	-1.4** (0.7)	-2.0%	(-2.6, -0.2)	0.05	75	71	0.1 (1.1)	0.2%	(-1.6, 1.9)	0.89	68	72	-2.6*** (1.0)	-3.7%	(-4.2, -1.0)	0.01
PY 3	69	70	-2.3*** (0.7)	-3.2%	(-3.5, -1.1)	0.00	73	70	-0.8 (1.1)	-1.1%	(-2.6, 1.0)	0.47	65	70	-3.5*** (1.0)	-5.1%	(-5.1, -1.8)	0.00
PY 4	54	55	-1.4* (0.8)	-2.6%	(-2.7, -0.2)	0.06	58	54	-0.2 (1.2)	-0.3%	(-2.1, 1.8)	0.87	51	55	-2.4** (1.0)	-4.5%	(-4.0, -0.8)	0.02
PY 5	54	55	-1.4* (0.8)	-2.5%	(-2.7, -0.1)	0.08	59	55	-0.4 (1.2)	-0.6%	(-2.4, 1.7)	0.76	51	55	-2.4** (1.0)	-4.5%	(-4.0, -0.7)	0.02
PY 1 through 5	64	65	-1.5** (0.6)	-2.4%	(-2.6, -0.5)	0.01	68	65	-0.4 (1.0)	-0.6%	(-2.0, 1.2)	0.69	60	65	-2.5*** (0.8)	-4.0%	(-3.8, -1.1)	0.00
Unweighted sample sizes																		
Number of practices	1,515	3,783					636	1,817					879	1,966				
Number of beneficiaries	1,896,880	4,507,499					847,208	2,257,322					1,053,634	2,261,852				

Source: Mathematica's analysis of Medicare claims data from January 2016 through December 2021.

Notes: Although this table indicates which estimates are statistically significant, when we interpret evidence, we combine evidence from the magnitude of the effect, the p-values, findings on related outcomes, subgroups, sensitivity tests, and other data sources about model implementation.

[a] We report the actual, unadjusted averages in the baseline period which are similar for the CPC+ and comparison groups because of matching. In the intervention periods, we computed the comparison group mean by subtracting the regression adjusted difference between the CPC+ and comparison means in each time period from the CPC+ mean in that same time period.

Table 5.J.5. ***(continued)***

[b] We regression-adjusted each impact estimate using a difference-in-differences analysis that reflects the difference of the average outcome for Medicare FFS beneficiaries attributed to CPC+ practices in the five years of CPC+ and the average outcome in the baseline year, relative to the same difference over time for Medicare FFS beneficiaries attributed to comparison practices, while controlling for beneficiary characteristics and practice fixed effects.

[c] We calculated percentage impacts relative to what the CPC+ mean would have been in PY 1 through PY 5 (separately and combined) in the absence of the intervention—that is, the unadjusted CPC+ mean minus the impact estimate.

*/**/*** Significantly different from zero at the 0.10/0.05/0.01 level, two-tailed test.

NA = not applicable because the difference-in-differences impact estimate cannot be calculated at baseline.

CC = complication or comorbidity; CPC+ = Comprehensive Primary Care Plus; FFS = fee-for-service; MCC = major complication or comorbidity; PY = Program Year; SE = standard error.

Table 5.J.6. Regression-adjusted means and estimated impact of CPC+ on types of surgical hospitalizations for attributed Medicare FFS beneficiaries over the five program years, Track 1

	Track 1—Overall						Track 1—SSP						Track 1—Non-SSP					
	CPC+ mean[a]	C mean[a]	Impact estimate[b] (SE)	Percentage impact[c]	90 percent confidence interval	*p*-Value	CPC+ mean[a]	C mean[a]	Impact estimate[b] (SE)	Percentage impact[c]	90 percent confidence interval	*p*-Value	CPC+ meana	C mean[a]	Impact estimate[b] (SE)	Percentage impact[c]	90 percent confidence interval	*p*-Value
Monthly Medicare expenditures by service category (per beneficiary per month)																		
Expenditures for acute surgical hospital admissions with an MCC																		
Baseline	$58	$58	NA	NA	NA	NA	$59	$59	NA	NA	NA	NA	$56	$57	NA	NA	NA	NA
PY 1	$57	$57	$0.7 ($1.0)	1.3%	(-$1.0, $2.4)	0.49	$58	$58	$0.6 ($1.4)	1.0%	(-$1.7, $2.9)	0.67	$56	$56	$0.8 ($1.5)	1.4%	(-$1.7, $3.3)	0.59
PY 2	$58	$60	-$0.7 ($1.0)	-1.2%	(-$2.3, $0.9)	0.47	$60	$61	-$0.9 ($1.3)	-1.4%	(-$3.1, $1.3)	0.51	$57	$58	-$0.5 ($1.5)	-0.9%	(-$2.9, $1.8)	0.71
PY 3	$61	$62	-$0.4 ($1.0)	-0.7%	(-$2.1, $1.2)	0.66	$61	$64	-$2.0 ($1.4)	-3.1%	(-$4.2, $0.3)	0.15	$60	$59	$1.2 ($1.5)	2.0%	(-$1.3, $3.7)	0.44
PY 4	$59	$61	-$1.1 ($1.0)	-1.8%	(-$2.8, $0.7)	0.31	$60	$63	-$2.6* ($1.4)	-4.2%	(-$5.0, -$0.3)	0.07	$59	$58	$0.7 ($1.5)	1.3%	(-$1.8, $3.3)	0.63
PY 5	$63	$66	-$2.3** ($1.1)	-3.5%	(-$4.1, -$0.5)	0.04	$65	$70	-$4.5*** ($1.5)	-6.4%	(-$7.0, -$2.0)	0.00	$61	$63	-$0.9 ($1.7)	-1.4%	(-$3.6, $1.9)	0.61
PY 1 through 5	$60	$61	-$0.8 ($0.8)	-1.2%	(-$2.1, $0.6)	0.35	$61	$63	-$1.8* ($1.1)	-2.9%	(-$3.6, $0.0)	0.10	$59	$59	$0.3 ($1.2)	0.4%	(-$1.8, $2.3)	0.83
Expenditures for acute surgical hospital admissions with a CC																		
Baseline	$30	$30	NA	NA	NA	NA	$31	$31	NA	NA	NA	NA	$29	$30	NA	NA	NA	NA
PY 1	$29	$29	$0.3 ($0.5)	1.2%	(-$0.4, $1.1)	0.48	$30	$30	-$0.2 ($0.6)	-0.8%	(-$1.3, $0.8)	0.71	$28	$28	$1.0 ($0.7)	3.5%	(-$0.2, $2.1)	0.18
PY 2	$30	$31	$0.0 ($0.5)	0.0%	(-$0.8, $0.8)	0.99	$32	$31	$0.3 ($0.7)	0.9%	(-$0.9, $1.4)	0.69	$29	$30	-$0.3 ($0.7)	-1.0%	(-$1.5, $0.9)	0.67
PY 3	$32	$33	-$0.1 ($0.5)	-0.4%	(-$0.9, $0.7)	0.79	$33	$33	-$0.1 ($0.7)	-0.3%	(-$1.2, $1.0)	0.89	$31	$32	-$0.2 ($0.7)	-0.6%	(-$1.4, $1.0)	0.80
PY 4	$30	$30	$0.3 ($0.5)	1.0%	(-$0.5, $1.1)	0.56	$30	$30	$0.0 ($0.7)	-0.1%	(-$1.2, $1.1)	0.97	$29	$29	$0.8 ($0.7)	2.8%	(-$0.4, $2.0)	0.28
PY 5	$31	$32	-$0.1 ($0.5)	-0.4%	(-$1.0, $0.8)	0.81	$32	$32	-$0.7 ($0.8)	-2.1%	(-$1.9, $0.6)	0.39	$30	$31	$0.3 ($0.8)	1.0%	(-$1.0, $1.6)	0.70
PY 1 through 5	$30	$31	$0.1 ($0.4)	0.2%	(-$0.6, $0.7)	0.86	$31	$32	-$0.1 ($0.5)	-0.4%	(-$1.0, $0.8)	0.81	$30	$30	$0.3 ($0.6)	1.0%	(-$0.6, $1.2)	0.61

Table 5.J.6. ***(continued)***

	Track 1—Overall						Track 1—SSP						Track 1—Non-SSP					
	CPC+ mean[a]	C mean[a]	Impact estimate[b] (SE)	Percentage impact[c]	90 percent confidence interval	*p*-Value	CPC+ mean[a]	C mean[a]	Impact estimate[b] (SE)	Percentage impact[c]	90 percent confidence interval	*p*-Value	CPC+ meana	C mean[a]	Impact estimate[b] (SE)	Percentage impact[c]	90 percent confidence interval	*p*-Value
Expenditures for acute surgical hospital admissions without an MCC or CC																		
Baseline	$60	$60	NA	NA	NA	NA	$62	$61	NA	NA	NA	NA	$58	$60	NA	NA	NA	NA
PY 1	$62	$61	$1.6* ($0.9)	2.6%	($0.0, $3.1)	0.09	$64	$62	$0.0 ($1.3)	0.1%	(-$2.1, $2.1)	0.98	$61	$59	$3.2** ($1.4)	5.7%	($1.0, $5.5)	0.02
PY 2	$60	$59	$1.1 ($0.9)	1.9%	(-$0.3, $2.5)	0.21	$62	$60	$0.3 ($1.2)	0.5%	(-$1.7, $2.3)	0.81	$58	$58	$2.0 ($1.3)	3.6%	(-$0.1, $4.1)	0.12
PY 3	$61	$61	$0.1 ($0.9)	0.1%	(-$1.4, $1.6)	0.94	$63	$62	-$0.3 ($1.3)	-0.5%	(-$2.4, $1.8)	0.81	$59	$59	$0.5 ($1.3)	0.9%	(-$1.6, $2.7)	0.69
PY 4	$50	$50	-$0.3 ($0.9)	-0.5%	(-$1.7, $1.2)	0.78	$51	$50	-$1.4 ($1.3)	-2.6%	(-$3.5, $0.8)	0.29	$48	$49	$1.0 ($1.3)	2.1%	(-$1.1, $3.1)	0.42
PY 5	$49	$47	$1.4 ($1.0)	3.1%	(-$0.2, $3.0)	0.14	$50	$49	$0.3 ($1.4)	0.6%	(-$1.9, $2.5)	0.82	$47	$46	$2.6* ($1.4)	5.8%	($0.3, $4.9)	0.07
PY 1 through 5	$56	$55	$0.8 ($0.7)	1.4%	(-$0.4, $2.0)	0.28	$58	$56	-$0.2 ($1.0)	-0.4%	(-$1.8, $1.4)	0.84	$54	$54	$1.8* ($1.0)	3.5%	($0.1, $3.6)	0.08
Annualized service use (per 1,000 beneficiaries per year)																		
Annualized acute surgical hospital admissions with an MCC																		
Baseline	22	22	NA	NA	NA	NA	22	22	NA	NA	NA	NA	22	22	NA	NA	NA	NA
PY 1	22	21	0.5 (0.3)	2.2%	(-0.1, 1.0)	0.14	22	21	0.4 (0.4)	1.9%	(-0.3, 1.1)	0.33	22	21	0.5 (0.5)	2.5%	(-0.3, 1.3)	0.28
PY 2	22	22	0.0 (0.3)	-0.2%	(-0.5, 0.5)	0.91	22	22	0.2 (0.4)	1.0%	(-0.4, 0.9)	0.59	22	22	-0.3 (0.5)	-1.4%	(-1.1, 0.5)	0.51
PY 3	22	22	0.1 (0.3)	0.4%	(-0.4, 0.6)	0.77	22	22	-0.3 (0.4)	-1.2%	(-1.0, 0.4)	0.51	23	21	0.5 (0.5)	2.3%	(-0.3, 1.3)	0.31
PY 4	21	20	0.0 (0.3)	-0.2%	(-0.6, 0.5)	0.92	21	21	-0.6 (0.4)	-2.8%	(-1.3, 0.1)	0.16	21	20	0.6 (0.5)	2.7%	(-0.3, 1.4)	0.27
PY 5	21	21	-0.3 (0.3)	-1.2%	(-0.8, 0.3)	0.44	22	22	-0.7 (0.4)	-3.2%	(-1.4, 0.0)	0.11	21	20	0.1 (0.5)	0.4%	(-0.8, 0.9)	0.88
PY 1 through 5	22	21	0.0 (0.3)	0.2%	(-0.4, 0.5)	0.86	22	22	-0.2 (0.3)	-0.8%	(-0.7, 0.4)	0.60	22	21	0.3 (0.4)	1.2%	(-0.4, 0.9)	0.52

Table 5.J.6. ***(continued)***

	Track 1—Overall						Track 1—SSP						Track 1—Non-SSP					
	CPC+ mean[a]	C mean[a]	Impact estimate[b] (SE)	Percentage impact[c]	90 percent confidence interval	*p*-Value	CPC+ mean[a]	C mean[a]	Impact estimate[b] (SE)	Percentage impact[c]	90 percent confidence interval	*p*-Value	CPC+ meana	C mean[a]	Impact estimate[b] (SE)	Percentage impact[c]	90 percent confidence interval	*p*-Value
Annualized acute surgical hospital admissions with a CC																		
Baseline	22	22	NA	NA	NA	NA	22	22	NA	NA	NA	NA	22	22	NA	NA	NA	NA
PY 1	21	21	0.3 (0.3)	1.4%	(-0.2, 0.8)	0.33	21	21	-0.1 (0.4)	-0.6%	(-0.8, 0.5)	0.74	21	20	0.8* (0.5)	3.7%	(0.0, 1.5)	0.10
PY 2	22	21	0.2 (0.3)	0.9%	(-0.3, 0.7)	0.54	22	21	0.4 (0.4)	1.7%	(-0.3, 1.1)	0.39	21	21	0.0 (0.5)	0.0%	(-0.8, 0.7)	0.99
PY 3	22	22	-0.1 (0.3)	-0.2%	(-0.6, 0.5)	0.87	23	22	-0.1 (0.4)	-0.3%	(-0.7, 0.6)	0.89	22	22	0.0 (0.5)	-0.2%	(-0.8, 0.7)	0.93
PY 4	20	19	0.3 (0.3)	1.7%	(-0.2, 0.8)	0.28	20	20	0.2 (0.4)	0.9%	(-0.5, 0.9)	0.66	20	19	0.5 (0.5)	2.8%	(-0.2, 1.3)	0.24
PY 5	20	20	0.1 (0.3)	0.6%	(-0.4, 0.6)	0.72	20	20	-0.1 (0.4)	-0.6%	(-0.9, 0.6)	0.77	20	19	0.4 (0.5)	1.9%	(-0.4, 1.1)	0.44
PY 1 through 5	21	20	0.2 (0.2)	0.8%	(-0.2, 0.6)	0.50	21	21	0.1 (0.3)	0.3%	(-0.5, 0.6)	0.87	21	20	0.3 (0.4)	1.5%	(-0.3, 0.9)	0.41
Annualized acute surgical hospital admissions without an MCC or CC																		
Baseline	46	44	NA	NA	NA	NA	46	44	NA	NA	NA	NA	46	44	NA	NA	NA	NA
PY 1	45	44	0.3 (0.5)	0.6%	(-0.5, 1.0)	0.57	45	44	-0.2 (0.6)	-0.4%	(-1.1, 0.8)	0.76	45	43	0.8 (0.7)	1.7%	(-0.4, 1.9)	0.30
PY 2	43	41	0.0 (0.5)	0.0%	(-0.8, 0.7)	0.98	42	41	-0.4 (0.6)	-1.0%	(-1.4, 0.5)	0.45	43	41	0.4 (0.7)	1.1%	(-0.7, 1.6)	0.53
PY 3	42	41	-0.1 (0.4)	-0.3%	(-0.8, 0.6)	0.80	42	41	-0.2 (0.6)	-0.5%	(-1.2, 0.7)	0.70	43	41	0.0 (0.7)	0.1%	(-1.1, 1.1)	0.97
PY 4	31	30	-1.2*** (0.4)	-3.7%	(-1.9, -0.4)	0.01	30	30	-1.6*** (0.6)	-4.9%	(-2.5, -0.6)	0.01	31	30	-0.8 (0.7)	-2.7%	(-2.0, 0.3)	0.21
PY 5	28	27	-0.5 (0.4)	-1.8%	(-1.2, 0.2)	0.27	28	28	-1.5** (0.6)	-5.0%	(-2.4, -0.5)	0.01	28	26	0.4 (0.7)	1.5%	(-0.7, 1.5)	0.56
PY 1 through 5	37	36	-0.3 (0.4)	-0.8%	(-0.9, 0.3)	0.41	37	36	-0.7 (0.5)	-1.9%	(-1.5, 0.0)	0.12	38	36	0.2 (0.6)	0.4%	(-0.7, 1.1)	0.77
Unweighted sample sizes																		
Number of practices	1,373	5,243					738	2,979					635	2,264				
Number of beneficiaries	1,549,585	5,347,499					798,817	3,129,830					753,337	2,233,041				

Source: Mathematica's analysis of Medicare claims data from January 2016 through December 2021.

Notes: Although this table indicates which estimates are statistically significant, when we interpret evidence, we combine evidence from the magnitude of the effect, the p-values, findings on related outcomes, subgroups, sensitivity tests, and other data sources about model implementation.

[a] We report the actual, unadjusted averages in the baseline period which are similar for the CPC+ and comparison groups because of matching. In the intervention periods, we computed the comparison group mean by subtracting the regression adjusted difference between the CPC+ and comparison means in each time period from the CPC+ mean in that same time period.

Table 5.J.6. ***(continued)***

[b] We regression-adjusted each impact estimate using a difference-in-differences analysis that reflects the difference of the average outcome for Medicare FFS beneficiaries attributed to CPC+ practices in the five years of CPC+ and the average outcome in the baseline year, relative to the same difference over time for Medicare FFS beneficiaries attributed to comparison practices, while controlling for beneficiary characteristics and practice fixed effects.

[c] We calculated percentage impacts relative to what the CPC+ mean would have been in PY 1 through PY 5 (separately and combined) in the absence of the intervention—that is, the unadjusted CPC+ mean minus the impact estimate.

*/**/*** Significantly different from zero at the 0.10/0.05/0.01 level, two-tailed test.

NA = not applicable, because the difference-in-differences impact estimate cannot be calculated at baseline.

CC = complication or comorbidity; CPC+ = Comprehensive Primary Care Plus; FFS = fee-for-service; MCC = major complication or comorbidity; PY = Program Year; SE = standard error.

Table 5.J.7. Regression-adjusted means and estimated impact of CPC+ on types of surgical hospitalizations for attributed Medicare FFS beneficiaries over the five program years, Track 2

	Track 2—Overall						Track 2—SSP						Track 2—Non-SSP					
	CPC+ mean[a]	C mean[a]	Impact estimate[b] (SE)	Percentage impact[c]	90 percent confidence interval	*p*-Value	CPC+ mean[a]	C mean[a]	Impact estimate[b] (SE)	Percentage impact[c]	90 percent confidence interval	*p*-Value	CPC+ mean[a]	C mean[a]	Impact estimate[b] (SE)	Percentage impact[c]	90 percent confidence interval	*p*-Value
Monthly Medicare expenditures by service category (per beneficiary per month)																		
Expenditures for acute surgical hospital admissions with an MCC																		
Baseline	$57	$57	NA	NA	NA	NA	$59	$59	NA	NA	NA	NA	$55	$56	NA	NA	NA	NA
PY 1	$57	$57	$1.0 ($1.0)	1.8%	(-$0.6, $2.7)	0.32	$58	$59	-$1.4 ($1.5)	-2.4%	(-$3.9, $1.1)	0.35	$57	$55	$2.9** ($1.3)	5.5%	($0.7, $5.1)	0.03
PY 2	$59	$59	$1.3 ($1.1)	2.2%	(-$0.5, $3.0)	0.24	$60	$62	-$2.6 ($1.7)	-4.2%	(-$5.4, $0.1)	0.12	$59	$56	$4.3*** ($1.4)	7.9%	($2.1, $6.6)	0.00
PY 3	$62	$63	-$0.2 ($1.1)	-0.2%	(-$1.9, $1.6)	0.89	$63	$65	-$2.4 ($1.6)	-3.7%	(-$5.1, $0.2)	0.13	$61	$61	$1.6 ($1.5)	2.7%	(-$0.8, $4.0)	0.27
PY 4	$59	$61	-$1.1 ($1.1)	-1.8%	(-$2.9, $0.8)	0.34	$61	$64	-$3.6** ($1.7)	-5.7%	(-$6.5, -$0.8)	0.03	$58	$59	$1.5 ($1.4)	2.7%	(-$0.8, $3.9)	0.28
PY 5	$64	$66	-$1.9 ($1.2)	-2.9%	(-$3.8, $0.0)	0.11	$65	$69	-$5.1*** ($1.8)	-7.3%	(-$8.1, -$2.1)	0.00	$63	$63	$1.3 ($1.5)	2.1%	(-$1.2, $3.8)	0.40
PY 1 through 5	$60	$61	-$0.2 ($0.9)	-0.3%	(-$1.6, $1.3)	0.85	$61	$64	-$3.0** ($1.3)	-4.7%	(-$5.2, -$0.9)	0.02	$60	$59	$2.4** ($1.1)	4.1%	($0.5, $4.2)	0.04
Expenditures for acute surgical hospital admissions with a CC																		
Baseline	$30	$31	NA	NA	NA	NA	$30	$31	NA	NA	NA	NA	$29	$30	NA	NA	NA	NA
PY 1	$29	$29	$0.7 ($0.5)	2.6%	(-$0.1, $1.6)	0.15	$29	$30	$0.8 ($0.7)	2.8%	(-$0.4, $2.0)	0.27	$29	$29	$0.7 ($0.7)	2.4%	(-$0.5, $1.8)	0.34
PY 2	$31	$31	$0.8 ($0.5)	2.7%	($0.0, $1.7)	0.12	$31	$32	$0.7 ($0.8)	2.4%	(-$0.6, $2.0)	0.36	$30	$30	$0.9 ($0.7)	3.0%	(-$0.3, $2.0)	0.20
PY 3	$32	$33	$0.2 ($0.5)	0.6%	(-$0.6, $1.0)	0.70	$33	$34	$0.7 ($0.8)	2.2%	(-$0.6, $2.1)	0.37	$31	$32	-$0.2 ($0.7)	-0.7%	(-$1.3, $0.9)	0.74
PY 4	$30	$30	$0.7 ($0.5)	2.5%	(-$0.1, $1.6)	0.16	$31	$31	$1.4* ($0.8)	4.8%	($0.1, $2.7)	0.07	$29	$30	$0.2 ($0.7)	0.5%	(-$1.0, $1.3)	0.82
PY 5	$31	$31	$1.0* ($0.5)	3.2%	($0.1, $1.9)	0.08	$32	$31	$1.9** ($0.9)	6.3%	($0.5, $3.3)	0.03	$31	$31	$0.4 ($0.7)	1.2%	(-$0.8, $1.5)	0.63
PY 1 through 5	$31	$31	$0.7 ($0.4)	2.3%	($0.0, $1.4)	0.10	$32	$32	$1.1* ($0.6)	3.6%	($0.1, $2.1)	0.08	$30	$31	$0.4 ($0.6)	1.2%	(-$0.5, $1.3)	0.51

Table 5.J.7. ***(continued)***

	Track 2—Overall						Track 2—SSP						Track 2—Non-SSP					
	CPC+ mean[a]	C mean[a]	Impact estimate[b] (SE)	Percentage impact[c]	90 percent confidence interval	*p*-Value	CPC+ mean[a]	C mean[a]	Impact estimate[b] (SE)	Percentage impact[c]	90 percent confidence interval	*p*-Value	CPC+ mean[a]	C mean[a]	Impact estimate[b] (SE)	Percentage impact[c]	90 percent confidence interval	*p*-Value
Expenditures for acute surgical hospital admissions without an MCC or CC																		
Baseline	$60	$60	NA	NA	NA	NA	$62	$60	NA	NA	NA	NA	$59	$60	NA	NA	NA	NA
PY 1	$63	$60	$2.4*** ($0.9)	3.9%	($0.9, $3.9)	0.01	$64	$62	$1.1 ($1.3)	1.8%	(-$1.1, $3.4)	0.39	$62	$59	$3.4*** ($1.2)	5.8%	($1.3, $5.4)	0.01
PY 2	$61	$59	$1.3 ($0.9)	2.2%	(-$0.2, $2.8)	0.15	$61	$60	$0.0 ($1.3)	0.0%	(-$2.2, $2.2)	1.00	$61	$59	$2.3* ($1.2)	4.0%	($0.3, $4.3)	0.06
PY 3	$61	$61	-$0.6 ($0.9)	-1.0%	(-$2.0, $0.8)	0.49	$62	$62	-$1.7 ($1.3)	-2.7%	(-$3.8, $0.4)	0.19	$60	$60	$0.3 ($1.1)	0.5%	(-$1.6, $2.1)	0.81
PY 4	$51	$49	$1.1 ($1.0)	2.3%	(-$0.5, $2.7)	0.25	$50	$50	-$0.9 ($1.5)	-1.8%	(-$3.3, $1.5)	0.55	$51	$48	$3.2*** ($1.2)	6.7%	($1.2, $5.2)	0.01
PY 5	$48	$46	$1.6* ($1.0)	3.5%	($0.0, $3.2)	0.09	$48	$47	-$0.2 ($1.4)	-0.4%	(-$2.5, $2.1)	0.89	$48	$46	$3.1** ($1.3)	6.8%	($0.9, $5.2)	0.02
PY 1 through 5	$56	$55	$1.1 ($0.7)	2.0%	($0.0, $2.3)	0.11	$57	$56	-$0.4 ($1.0)	-0.6%	(-$2.0, $1.3)	0.72	$56	$54	$2.4** ($0.9)	4.4%	($0.8, $3.9)	0.01
Annualized service use (per 1,000 beneficiaries per year)																		
Annualized acute surgical hospital admissions with an MCC																		
Baseline	22	22	NA	NA	NA	NA	23	22	NA	NA	NA	NA	22	21	NA	NA	NA	NA
PY 1	22	21	0.1 (0.3)	0.3%	(-0.5, 0.6)	0.83	22	22	-1.2** (0.5)	-5.0%	(-1.9, -0.4)	0.02	22	21	1.0** (0.4)	4.9%	(0.4, 1.7)	0.01
PY 2	22	22	0.1 (0.3)	0.4%	(-0.4, 0.6)	0.77	22	22	-1.1** (0.5)	-4.6%	(-1.9, -0.3)	0.03	22	21	1.0** (0.4)	4.7%	(0.3, 1.7)	0.02
PY 3	23	22	-0.2 (0.3)	-0.8%	(-0.7, 0.4)	0.58	23	23	-1.0** (0.5)	-4.4%	(-1.9, -0.2)	0.04	22	22	0.5 (0.5)	2.1%	(-0.3, 1.2)	0.31
PY 4	21	21	-0.5 (0.3)	-2.4%	(-1.0, 0.0)	0.12	21	21	-1.3*** (0.5)	-5.9%	(-2.1, -0.5)	0.01	21	20	0.2 (0.4)	1.2%	(-0.5, 0.9)	0.58
PY 5	21	21	-0.5 (0.3)	-2.5%	(-1.1, 0.0)	0.12	21	22	-1.8*** (0.5)	-7.7%	(-2.7, -0.9)	0.00	21	20	0.6 (0.4)	3.1%	(-0.1, 1.4)	0.16
PY 1 through 5	22	21	-0.2 (0.3)	-1.0%	(-0.7, 0.2)	0.43	22	22	-1.3*** (0.4)	-5.5%	(-1.9, -0.6)	0.00	22	21	0.7* (0.4)	3.2%	(0.1, 1.3)	0.05

Table 5.J.7. ***(continued)***

	Track 2—Overall						Track 2—SSP						Track 2—Non-SSP					
	CPC+ mean[a]	C mean[a]	Impact estimate[b] (SE)	Percentage impact[c]	90 percent confidence interval	*p*-Value	CPC+ mean[a]	C mean[a]	Impact estimate[b] (SE)	Percentage impact[c]	90 percent confidence interval	*p*-Value	CPC+ mean[a]	C mean[a]	Impact estimate[b] (SE)	Percentage impact[c]	90 percent confidence interval	*p*-Value
Annualized acute surgical hospital admissions with a CC																		
Baseline	22	22	NA	NA	NA	NA	22	22	NA	NA	NA	NA	22	22	NA	NA	NA	NA
PY 1	21	21	0.2 (0.3)	1.2%	(-0.3, 0.8)	0.46	21	21	0.4 (0.5)	2.1%	(-0.3, 1.2)	0.34	21	21	0.1 (0.5)	0.4%	(-0.7, 0.9)	0.84
PY 2	22	21	0.3 (0.3)	1.4%	(-0.2, 0.8)	0.35	22	22	0.3 (0.5)	1.6%	(-0.4, 1.1)	0.45	21	21	0.3 (0.4)	1.3%	(-0.4, 1.0)	0.54
PY 3	22	22	0.1 (0.3)	0.6%	(-0.4, 0.7)	0.70	23	22	0.5 (0.5)	2.2%	(-0.4, 1.3)	0.34	22	22	-0.2 (0.4)	-0.7%	(-0.9, 0.6)	0.72
PY 4	20	20	0.3 (0.3)	1.3%	(-0.3, 0.8)	0.43	20	20	0.8* (0.5)	3.9%	(0.0, 1.5)	0.10	19	19	-0.1 (0.4)	-0.7%	(-0.9, 0.6)	0.76
PY 5	20	19	0.8** (0.3)	3.9%	(0.2, 1.3)	0.02	21	19	1.3*** (0.5)	6.7%	(0.5, 2.1)	0.01	20	19	0.4 (0.4)	2.1%	(-0.3, 1.1)	0.34
PY 1 through 5	21	21	0.3 (0.3)	1.6%	(-0.1, 0.8)	0.21	21	21	0.7* (0.4)	3.2%	(0.0, 1.3)	0.09	21	20	0.1 (0.4)	0.5%	(-0.5, 0.7)	0.80
Annualized acute surgical hospital admissions without an MCC or CC																		
Baseline	46	44	NA	NA	NA	NA	46	44	NA	NA	NA	NA	46	44	NA	NA	NA	NA
PY 1	45	43	0.4 (0.4)	1.0%	(-0.3, 1.2)	0.33	46	44	0.2 (0.6)	0.4%	(-0.8, 1.2)	0.77	45	43	0.6 (0.6)	1.4%	(-0.4, 1.7)	0.31
PY 2	43	41	0.5 (0.4)	1.1%	(-0.3, 1.2)	0.31	43	41	0.4 (0.6)	1.0%	(-0.7, 1.5)	0.53	43	41	0.5 (0.6)	1.1%	(-0.5, 1.5)	0.43
PY 3	43	41	0.0 (0.4)	0.0%	(-0.7, 0.7)	0.97	43	41	0.6 (0.7)	1.4%	(-0.5, 1.6)	0.37	42	41	-0.4 (0.6)	-1.0%	(-1.4, 0.5)	0.47
PY 4	30	29	-0.6 (0.4)	-2.1%	(-1.4, 0.1)	0.14	30	30	-1.2* (0.7)	-4.0%	(-2.3, -0.2)	0.06	31	30	-0.3 (0.6)	-1.1%	(-1.3, 0.6)	0.55
PY 5	27	26	-0.8* (0.4)	-2.8%	(-1.5, -0.1)	0.07	27	26	-1.1* (0.6)	-3.8%	(-2.1, 0.0)	0.10	27	26	-0.6 (0.6)	-2.3%	(-1.6, 0.4)	0.31
PY 1 through 5	37	36	-0.1 (0.4)	-0.3%	(-0.7, 0.5)	0.78	37	36	-0.2 (0.5)	-0.5%	(-1.0, 0.6)	0.70	37	36	0.0 (0.5)	-0.1%	(-0.9, 0.8)	0.92
Unweighted sample sizes																		
Number of practices	1,515	3,783					636	1,817					879	1,966				
Number of beneficiaries	1,896,880	4,507,499					847,208	2,257,322					1,053,634	2,261,852				

Source: Mathematica's analysis of Medicare claims data from January 2016 through December 2021.

Notes: Although this table indicates which estimates are statistically significant, when we interpret evidence, we combine evidence from the magnitude of the effect, the p-values, findings on related outcomes, subgroups, sensitivity tests, and other data sources about model implementation.

[a] We report the actual, unadjusted averages in the baseline period which are similar for the CPC+ and comparison groups because of matching. In the intervention periods, we computed the comparison group mean by subtracting the regression adjusted difference between the CPC+ and comparison means in each time period from the CPC+ mean in that same time period.

Table 5.J.7. ***(continued)***

[b] We regression-adjusted each impact estimate using a difference-in-differences analysis that reflects the difference of the average outcome for Medicare FFS beneficiaries attributed to CPC+ practices in the five years of CPC+ and the average outcome in the baseline year, relative to the same difference over time for Medicare FFS beneficiaries attributed to comparison practices, while controlling for beneficiary characteristics and practice fixed effects.

[c] We calculated percentage impacts relative to what the CPC+ mean would have been in PY 1 through PY 5 (separately and combined) in the absence of the intervention—that is, the unadjusted CPC+ mean minus the impact estimate.

*/**/*** Significantly different from zero at the 0.10/0.05/0.01 level, two-tailed test.

NA = not applicable, because the difference-in-differences impact estimate cannot be calculated at baseline.

CC = complication or comorbidity; CPC+ = Comprehensive Primary Care Plus; FFS = fee-for-service; MCC = major complication or comorbidity; PY = Program Year; SE = standard error.

D. Impact on hospitalizations by practice subgroup

In the following section, we describe key findings from the practice subgroup analyses. We focus on subgroups with multiple statistical differences in the impact estimates between subgroups (suggesting a pattern in differences across groups) and statistically significant impacts in at least one of the subgroups.

Impact estimates differed meaningfully between practices owned by a hospital or health system and independent practices across multiple outcomes; reductions were concentrated among independent practices in both tracks. Over the five years of the model, CPC+ reduced all-cause acute hospital admissions for independent practices but had no effect for the hospital- or system-owned group (Appendix 5.A). When studying the more detailed hospital outcomes, we see a similar pattern:

- In Track 1, there were reductions in surgical admissions for independent practices (1.8 percent), but there was no impact for the hospital- or system-owned group. The differences in estimates were statistically significant ($p < 0.01$) and driven by differences within non-SSP practices (Table 5.J.8). There was also a statistical difference between these two groups in the estimate for expenditures for surgical admissions, though neither practice type had statistically significant impacts.
- For Track 2, over the five years of the model, CPC+ reduced expenditures for acute medical admissions among independent practices by $6.6 PBPM (4.5 percent) and reduced the number of medical admissions for independent practices by 3.7 percent but did not affect outcomes for the group of practices that were hospital- or system-owned. These differences across practice groups existed for both SSP and non-SSP practices (Table 5.J.9).

When stratifying by urbanicity of the practice's county, findings indicate statistical differences across practice groups, with urban practices having more favorable impacts.

- In Track 1, urban practices reduced expenditures for medical hospital admissions (4.9 per 1,000 beneficiaries per year) versus no impacts for other two urbanicity categories (rural and suburban; Table 5.J.8)
- In Track 2, there was an increase in expenditures for surgical hospital admission among rural (8.5 percent) and suburban (4.1 percent) practices, with no impact for urban practices. There were also statistical differences in the impact of CPC+ on the number of surgical hospitalizations. Rural practices showed an increase in surgical admissions (5.3 percent). However, there were no impacts for suburban or urban practices (Table 5. J.9).

In Track 2, there were meaningful differences in the impact of CPC+ on medical hospitalizations for practices that participated in prior primary care transformation initiatives versus those that did not, with more favorable impacts for practices with transformation experience.

- Over the five program years, Track 2 practices with prior transformation experience had reductions of 3.9 admissions per 1,000 beneficiaries per year (2 percent of the baseline value) for acute medical hospital admissions. Practices without transformation experience had no impacts. Similarly, there were reductions in expenditures for medical hospital admissions for practices with prior transformation experience (3.6 percent) versus no impact for practices without transformation experience. For both outcomes, these differences were driven by differences within SSP practices (Table 5.J.9).

Table 5.J.8. Estimates of five-year impacts of CPC+ on surgical and medical hospitalizations for Track 1, by baseline practice characteristics

		Track 1– Overall				Track 1–SSP				Track 1–-Non-SSP			
	Practice subgroup definition, based on baseline characteristics	Number (percentage) of CPC+ beneficiaries in subgroup at baseline	Impact estimate (standard error)	Percentage impact	*p*-Value for difference in impact estimates between subgroups[a]	Number (percentage) of CPC+ beneficiaries in subgroup at baseline	Impact estimate (standard error)	Percentage impact	*p*-Value for difference in impact estimates between subgroups[a]	Number (percentage) of CPC+ beneficiaries in subgroup at baseline	Impact estimate (standard error)	Percentage impact	*p*-Value for difference in impact estimates between subgroups[a]
Monthly Medicare expenditures for surgical hospital admissions (per beneficiary per month)													
Whether practice participated in prior primary care transformation initiatives (recognized as a medical home or participated in MAPCP or CPC Classic)													
PY 1 through PY 5	Yes	468,487 (53.6%)	$0.6 ($1.7)	0.4%		214,075 (47.7%)	-$1.7 ($2.5)	-1.1%		254,262 (59.8%)	$2.5 ($2.3)	1.8%	
	No	405,383 (46.4%)	-$0.6 ($1.8)	-0.4%	0.52	234,948 (52.3%)	-$2.4 ($2.3)	-1.6%	0.34	170,586 (40.2%)	$1.6 ($2.8)	1.1%	0.93
Large and medium, versus small practice based on number of primary care practitioners													
PY 1 through PY 5	Large (6+ primary care practitioners)	404,456 (46.3%)	-$1.3 ($1.8)	-0.9%		189,229 (42.1%)	-$4.2 ($2.7)	-2.8%		215,122 (50.6%)	$0.8 ($2.4)	0.5%	
	Medium (3–5 primary care practitioners)	282,380 (32.3%)	$1.2 ($2.2)	0.9%		156,338 (34.8%)	-$2.1 ($2.7)	-1.4%		126,106 (29.7%)	$5.6 ($3.5)	4.1%	
	Small (1–2 primary care practitioners)	187,034 (21.4%)	$1.3 ($2.6)	0.8%	0.27	103,455 (23.0%)	$1.9 ($3.5)	1.3%	0.37	83,621 (19.7%)	$0.4 ($3.9)	0.3%	0.41
Whether hospital- or system-owned													
PY 1 through PY 5	Hospital-or system-owned	474,606 (54.3%)	$2.1 ($1.7)	1.5%		250,558 (55.8%)	-$2.1 ($2.2)	-1.4%		224,086 (52.7%)	$6.4** ($2.6)	4.6%	
	Independent	399,264 (45.7%)	-$2.4 ($1.8)	-1.6%	0.06	198,464 (44.2%)	-$2.1 ($2.6)	-1.4%	0.72	200,762 (47.3%)	-$2.6 ($2.5)	-1.8%	0.01
Whether the practice shared a TIN with another primary care practice													
PY 1 through PY 5	Shared a TIN with another primary care practice	684,507 (78.3%)	$0.7 ($1.4)	0.5%		366,843 (81.7%)	-$2.1 ($1.8)	-1.4%		317,749 (74.8%)	$3.7* ($2.0)	2.7%	
	Did not share a TIN with another primary care practice	189,364 (21.7%)	-$1.8 ($2.5)	-1.2%	0.33	82,179 (18.3%)	-$2.1 ($3.7)	-1.3%	0.72	107,099 (25.2%)	-$1.5 ($3.3)	-1.0%	0.23
Practice type: multi-specialty versus primary care only													
PY 1 through PY 5	Multi-specialty	170,691 (19.5%)	$2.0 ($3.0)	1.4%		76,547 (17.0%)	-$3.1 ($4.5)	-2.0%		94,082 (22.1%)	$5.2 ($4.1)	3.8%	
	Primary care only	703,179 (80.5%)	-$0.4 ($1.3)	-0.3%	0.22	372,475 (83.0%)	-$1.9 ($1.8)	-1.2%	0.83	330,766 (77.9%)	$1.2 ($2.0)	0.9%	0.13
Urbanicity of practice's county: rural or suburban versus urban													
PY 1 through PY 5	Rural	89,834 (10.3%)	-$2.7 ($3.7)	-1.9%		22,327 (5.0%)	-$2.1 ($7.2)	-1.4%		67,372 (15.9%)	-$2.0 ($4.3)	-1.5%	
	Suburban	156,799 (17.9%)	$3.6 ($2.9)	2.6%		74,982 (16.7%)	$6.9* ($3.7)	4.9%		81,785 (19.3%)	$0.3 ($4.3)	0.2%	
	Urban	627,237 (71.8%)	-$0.4 ($1.5)	-0.3%	0.31	351,712 (78.3%)	-$4.0 ($2.0)	-2.6%	0.03	275,691 (64.9%)	$3.7* ($2.2)	2.6%	0.24

Table 5.J.8. ***(continued)***

		Track 1– Overall				Track 1–SSP				Track 1–-Non-SSP			
	Practice subgroup definition, based on baseline characteristics	Number (percentage) of CPC+ beneficiaries in subgroup at baseline	Impact estimate (standard error)	Percentage impact	*p*-Value for difference in impact estimates between subgroups[a]	Number (percentage) of CPC+ beneficiaries in subgroup at baseline	Impact estimate (standard error)	Percentage impact	*p*-Value for difference in impact estimates between subgroups[a]	Number (percentage) of CPC+ beneficiaries in subgroup at baseline	Impact estimate (standard error)	Percentage impact	*p*-Value for difference in impact estimates between subgroups[a]
Monthly Medicare expenditures for medical hospital admissions (per beneficiary per month)													
Whether practice participated in prior primary care transformation initiatives (recognized as a medical home or participated in MAPCP or CPC Classic)													
PY 1 through PY 5	Yes	468,487 (53.6%)	-$2.6* ($1.4)	-1.9%		214,075 (47.7%)	-$2.4 ($1.8)	-1.7%		254,262 (59.8%)	-$2.8 ($2.0)	-2.1%	
	No	405,383 (46.4%)	-$3.8*** ($1.3)	-2.6%	0.57	234,948 (52.3%)	-$6.9*** ($1.8)	-4.4%	0.03	170,586 (40.2%)	-$0.3 ($2.0)	-0.2%	0.22
Large and medium, versus small practice based on number of primary care practitioners													
PY 1 through PY 5	Large (6+ primary care practitioners)	404,456 (46.3%)	-$3.7** ($1.5)	-2.7%		189,229 (42.1%)	-$7.3*** ($2.1)	-5.0%		215,122 (50.6%)	-$1.0 ($2.2)	-0.7%	
	Medium (3–5 primary care practitioners)	282,380 (32.3%)	-$0.9 ($1.7)	-0.6%		156,338 (34.8%)	-$0.5 ($2.0)	-0.4%		126,106 (29.7%)	-$1.5 ($2.7)	-1.0%	
	Small (1–2 primary care practitioners)	187,034 (21.4%)	-$5.5*** ($2.0)	-3.5%	0.19	103,455 (23.0%)	-$6.5** ($2.7)	-4.0%	0.16	83,621 (19.7%)	-$4.4 ($2.9)	-2.9%	0.53
Whether hospital- or system-owned													
PY 1 through PY 5	Hospital-or system-owned	474,606 (54.3%)	-$2.9** ($1.4)	-2.0%		250,558 (55.8%)	-$4.0** ($1.7)	-2.7%		224,086 (52.7%)	-$1.9 ($2.2)	-1.3%	
	Independent	399,264 (45.7%)	-$3.6*** ($1.4)	-2.5%	0.96	198,464 (44.2%)	-$5.7*** ($2.0)	-3.7%	0.85	200,762 (47.3%)	-$1.7 ($1.9)	-1.3%	0.90
Whether the practice shared a TIN with another primary care practice													
PY 1 through PY 5	Shared a TIN with another primary care practice	684,507 (78.3%)	-$2.7** ($1.1)	-1.9%		366,843 (81.7%)	-$3.6*** ($1.4)	-2.5%		317,749 (74.8%)	-$1.7 ($1.7)	-1.3%	
	Did not share a TIN with another primary care practice	189,364 (21.7%)	-$4.8** ($1.9)	-3.3%	0.39	82,179 (18.3%)	-$9.4*** ($3.1)	-6.1%	0.09	107,099 (25.2%)	-$1.5 ($2.4)	-1.1%	0.78
Practice type: multi-specialty versus primary care only													
PY 1 through PY 5	Multi-specialty	170,691 (19.5%)	-$2.7 ($2.4)	-1.9%		76,547 (17.0%)	-$8.8*** ($2.5)	-5.9%		94,082 (22.1%)	$2.0 ($3.9)	1.6%	
	Primary care only	703,179 (80.5%)	-$3.3*** ($1.1)	-2.3%	1.00	372,475 (83.0%)	-$3.9*** ($1.5)	-2.6%	0.22	330,766 (77.9%)	-$2.9* ($1.5)	-2.0%	0.58
Urbanicity of practice's county: rural or suburban versus urban													
PY 1 through PY 5	Rural	89,834 (10.3%)	$3.2 ($3.3)	2.5%		22,327 (5.0%)	-$2.7 ($6.8)	-2.0%		67,372 (15.9%)	$4.9 ($3.7)	3.9%	
	Suburban	156,799 (17.9%)	-$0.1 ($2.5)	-0.1%		74,982 (16.7%)	$1.3 ($3.2)	0.9%		81,785 (19.3%)	-$1.5 ($3.7)	-1.1%	
	Urban	627,237 (71.8%)	-$4.9*** ($1.1)	-3.3%	0.02	351,712 (78.3%)	-$6.1*** ($1.4)	-4.1%	0.13	275,691 (64.9%)	-$3.5** ($1.8)	-2.4%	0.15

Table 5.J.8. ***(continued)***

		Track 1– Overall				Track 1–SSP				Track 1–-Non-SSP			
	Practice subgroup definition, based on baseline characteristics	Number (percentage) of CPC+ beneficiaries in subgroup at baseline	Impact estimate (standard error)	Percentage impact	*p*-Value for difference in impact estimates between subgroups[a]	Number (percentage) of CPC+ beneficiaries in subgroup at baseline	Impact estimate (standard error)	Percentage impact	*p*-Value for difference in impact estimates between subgroups[a]	Number (percentage) of CPC+ beneficiaries in subgroup at baseline	Impact estimate (standard error)	Percentage impact	*p*-Value for difference in impact estimates between subgroups[a]
Annualized acute surgical hospital admissions (per 1,000 beneficiaries per year)													
Whether practice participated in prior primary care transformation initiatives (recognized as a medical home or participated in MAPCP or CPC Classic)													
PY 1 through PY 5	Yes	468,487 (53.6%)	0.3 (0.8)	0.3%		214,075 (47.7%)	-0.1 (1.0)	-0.2%		254,262 (59.8%)	0.5 (1.1)	0.6%	
	No	405,383 (46.4%)	-0.5 (0.8)	-0.6%	0.55	234,948 (52.3%)	-1.5 (1.0)	-1.8%	0.16	170,586 (40.2%)	0.7 (1.3)	0.9%	0.59
Large and medium, versus small practice based on number of primary care practitioners													
PY 1 through PY 5	Large (6+ primary care practitioners)	404,456 (46.3%)	-0.5 (0.8)	-0.6%		189,229 (42.1%)	-1.3 (1.1)	-1.6%		215,122 (50.6%)	0.1 (1.2)	0.1%	
	Medium (3–5 primary care practitioners)	282,380 (32.3%)	0.3 (1.0)	0.4%		156,338 (34.8%)	-0.6 (1.2)	-0.7%		126,106 (29.7%)	1.4 (1.7)	1.8%	
	Small (1–2 primary care practitioners)	187,034 (21.4%)	0.0 (1.1)	0.1%	0.48	103,455 (23.0%)	-0.3 (1.4)	-0.4%	0.76	83,621 (19.7%)	0.5 (1.7)	0.7%	0.53
Whether hospital- or system-owned													
PY 1 through PY 5	Hospital-or system-owned	474,606 (54.3%)	1.0 (0.8)	1.2%		250,558 (55.8%)	-0.8 (0.9)	-1.0%		224,086 (52.7%)	2.8** (1.2)	3.5%	
	Independent	399,264 (45.7%)	-1.4* (0.8)	-1.8%	0.03	198,464 (44.2%)	-0.8 (1.1)	-1.1%	0.67	200,762 (47.3%)	-1.9* (1.1)	-2.4%	0.00
Whether the practice shared a TIN with another primary care practice													
PY 1 through PY 5	Shared a TIN with another primary care practice	684,507 (78.3%)	0.4 (0.6)	0.5%		366,843 (81.7%)	-0.4 (0.7)	-0.4%		317,749 (74.8%)	1.1 (1.0)	1.4%	
	Did not share a TIN with another primary care practice	189,364 (21.7%)	-1.8* (1.1)	-2.2%	0.07	82,179 (18.3%)	-3.0* (1.5)	-3.6%	0.09	107,099 (25.2%)	-0.9 (1.5)	-1.2%	0.24
Practice type: multi-specialty versus primary care only													
PY 1 through PY 5	Multi-specialty	170,691 (19.5%)	1.0 (1.3)	1.3%		76,547 (17.0%)	-0.7 (1.6)	-0.9%		94,082 (22.1%)	2.3 (2.1)	3.1%	
	Primary care only	703,179 (80.5%)	-0.4 (0.6)	-0.5%	0.24	372,475 (83.0%)	-0.9 (0.8)	-1.1%	0.78	330,766 (77.9%)	0.1 (0.9)	0.1%	0.20
Urbanicity of practice's county: rural or suburban versus urban													
PY 1 through PY 5	Rural	89,834 (10.3%)	0.5 (1.7)	0.7%		22,327 (5.0%)	-1.0 (3.1)	-1.3%		67,372 (15.9%)	1.3 (2.1)	1.7%	
	Suburban	156,799 (17.9%)	0.9 (1.4)	1.1%		74,982 (16.7%)	1.4 (1.7)	1.7%		81,785 (19.3%)	0.1 (2.2)	0.2%	
	Urban	627,237 (71.8%)	-0.4 (0.6)	-0.5%	0.70	351,712 (78.3%)	-1.3* (0.8)	-1.6%	0.24	275,691 (64.9%)	0.5 (1.0)	0.6%	0.99

Table 5.J.8. ***(continued)***

		Track 1– Overall				Track 1–SSP				Track 1–-Non-SSP			
	Practice subgroup definition, based on baseline characteristics	Number (percentage) of CPC+ beneficiaries in subgroup at baseline	Impact estimate (standard error)	Percentage impact	*p*-Value for difference in impact estimates between subgroups[a]	Number (percentage) of CPC+ beneficiaries in subgroup at baseline	Impact estimate (standard error)	Percentage impact	*p*-Value for difference in impact estimates between subgroups[a]	Number (percentage) of CPC+ beneficiaries in subgroup at baseline	Impact estimate (standard error)	Percentage impact	*p*-Value for difference in impact estimates between subgroups[a]
Annualized acute medical hospital admissions (per 1,000 beneficiaries per year)													
Whether practice participated in prior primary care transformation initiatives (recognized as a medical home or participated in MAPCP or CPC Classic)													
PY 1 through PY 5	Yes	468,487 (53.6%)	-2.6 (1.7)	-1.4%		214,075 (47.7%)	-0.9 (2.2)	-0.5%		254,262 (59.8%)	-4.1* (2.4)	-2.2%	
	No	405,383 (46.4%)	-2.3 (1.7)	-1.2%	0.85	234,948 (52.3%)	-6.2*** (2.2)	-3.1%	0.09	170,586 (40.2%)	3.0 (2.5)	1.6%	0.07
Large and medium, versus small practice based on number of primary care practitioners													
PY 1 through PY 5	Large (6+ primary care practitioners)	404,456 (46.3%)	-4.7*** (1.8)	-2.5%		189,229 (42.1%)	-5.1** (2.5)	-2.7%		215,122 (50.6%)	-4.6* (2.5)	-2.4%	
	Medium (3–5 primary care practitioners)	282,380 (32.3%)	1.8 (2.2)	0.9%		156,338 (34.8%)	-0.2 (2.5)	-0.1%		126,106 (29.7%)	4.4 (3.6)	2.4%	
	Small (1–2 primary care practitioners)	187,034 (21.4%)	-4.2* (2.4)	-2.1%	0.04	103,455 (23.0%)	-6.3** (3.2)	-3.1%	0.37	83,621 (19.7%)	-1.3 (3.7)	-0.7%	0.14
Whether hospital- or system-owned													
PY 1 through PY 5	Hospital-or system-owned	474,606 (54.3%)	-2.4 (1.6)	-1.3%		250,558 (55.8%)	-3.5* (2.0)	-1.8%		224,086 (52.7%)	-1.2 (2.6)	-0.6%	
	Independent	399,264 (45.7%)	-2.6 (1.8)	-1.4%	0.95	198,464 (44.2%)	-3.9 (2.4)	-2.0%	0.74	200,762 (47.3%)	-1.3 (2.5)	-0.7%	0.73
Whether the practice shared a TIN with another primary care practice													
PY 1 through PY 5	Shared a TIN with another primary care practice	684,507 (78.3%)	-2.4* (1.4)	-1.3%		366,843 (81.7%)	-3.3* (1.7)	-1.7%		317,749 (74.8%)	-1.0 (2.1)	-0.5%	
	Did not share a TIN with another primary care practice	189,364 (21.7%)	-2.4 (2.4)	-1.3%	0.99	82,179 (18.3%)	-5.2 (3.7)	-2.6%	0.80	107,099 (25.2%)	-0.7 (3.0)	-0.4%	0.85
Practice type: multi-specialty versus primary care only													
PY 1 through PY 5	Multi-specialty	170,691 (19.5%)	0.0 (3.1)	0.0%		76,547 (17.0%)	-7.9** (3.6)	-4.1%		94,082 (22.1%)	6.6 (4.8)	4.0%	
	Primary care only	703,179 (80.5%)	-3.1** (1.3)	-1.6%	0.29	372,475 (83.0%)	-2.8 (1.7)	-1.4%	0.19	330,766 (77.9%)	-3.5* (1.8)	-1.8%	0.04
Urbanicity of practice's county: rural or suburban versus urban													
PY 1 through PY 5	Rural	89,834 (10.3%)	4.3 (4.3)	2.6%		22,327 (5.0%)	-2.7 (8.6)	-1.6%		67,372 (15.9%)	6.9 (5.0)	4.3%	
	Suburban	156,799 (17.9%)	-2.9 (3.3)	-1.5%		74,982 (16.7%)	-3.7 (4.1)	-1.8%		81,785 (19.3%)	-2.8 (4.8)	-1.6%	
	Urban	627,237 (71.8%)	-3.4** (1.3)	-1.7%	0.28	351,712 (78.3%)	-3.7** (1.7)	-1.9%	1.00	275,691 (64.9%)	-2.8 (2.1)	-1.4%	0.40

Source: Mathematica's analysis of Medicare claims data from January 2016 through December 2021.

Note: The estimates (and standard errors) in the impact estimate column show the separate subgroup-specific impacts over the first four years of CPC+ for each practice characteristic listed in the table.

Table 5.J.8. ***(continued)***

[a] The *p*-values in the last column represent the results from tests for statistically significant differences in impact estimates between the subgroups based on the baseline practice characteristic (using a t-test for subgroups with two categories and an F-test for subgroups with more than two categories).

*/**/***Within-subgroup estimate significantly different from zero at the 0.10/0.05/0.01 level, two-tailed test.

CPC = Comprehensive Primary Care; MAPCP = Multi-Payer Advanced Primary Care Practice; PY = Program Year; SSP = Medicare Shared Savings Program; TIN = Tax Identification Number.

Table 5.J.9. Estimates of five-year impacts of CPC+ on surgical and medical hospitalizations for Track 2, by baseline practice characteristics

		Track 2– Overall				Track 2–SSP				Track 2–Non-SSP			
	Practice subgroup definition, based on baseline characteristics	Number (percentage) of CPC+ beneficiaries in subgroup at baseline	Impact estimate (standard error)	Percentage impact	*p*-Value for difference in impact estimates between subgroups[a]	Number (percentage) of CPC+ beneficiaries in subgroup at baseline	Impact estimate (standard error)	Percentage impact	*p*-Value for difference in impact estimates between subgroups[a]	Number (percentage) of CPC+ beneficiaries in subgroup at baseline	Impact estimate (standard error)	Percentage impact	*p*-Value for difference in impact estimates between subgroups[a]
Monthly Medicare expenditures for surgical hospital admissions (per beneficiary per month)													
Whether practice participated in prior primary care transformation initiatives (recognized as a medical home or participated in MAPCP or CPC Classic)													
PY 1 through PY 5	Yes	865,798 (81.2%)	$0.6 ($1.5)	0.4%		385,875 (81.8%)	-$4.4** ($2.2)	-2.9%		479,947 (80.6%)	$5.5*** ($2.0)	3.9%	
	No	201,028 (18.8%)	$5.7** ($2.5)	3.9%	0.14	85,762 (18.2%)	$7.8* ($4.3)	5.2%	0.03	115,242 (19.4%)	$4.3 ($3.1)	3.0%	0.68
Large and medium, versus small practice based on number of primary care practitioners													
PY 1 through PY 5	Large (6+ primary care practitioners)	589,224 (55.2%)	$0.7 ($1.7)	0.4%		279,067 (59.2%)	-$3.2 ($2.7)	-2.1%		310,301 (52.1%)	$4.6** ($2.2)	3.3%	
	Medium (3–5 primary care practitioners)	340,406 (31.9%)	$1.5 ($2.3)	1.0%		134,103 (28.4%)	-$2.3 ($3.2)	-1.6%		206,177 (34.6%)	$5.2 ($3.2)	3.8%	
	Small (1–2 primary care practitioners)	137,196 (12.9%)	$5.4* ($3.3)	3.8%	0.61	58,467 (12.4%)	$2.7 ($4.7)	1.8%	0.95	78,712 (13.2%)	$7.5* ($4.4)	5.4%	0.59
Whether hospital- or system-owned													
PY 1 through PY 5	Hospital-or system-owned	619,957 (58.1%)	$2.5 ($1.7)	1.7%		289,350 (61.4%)	-$2.9 ($2.5)	-1.9%		330,724 (55.6%)	$7.6*** ($2.3)	5.5%	
	Independent	446,869 (41.9%)	$0.2 ($2.0)	0.2%	0.22	182,287 (38.6%)	-$1.0 ($3.1)	-0.7%	0.85	264,465 (44.4%)	$2.2 ($2.6)	1.6%	0.08
Whether the practice shared a TIN with another primary care practice													
PY 1 through PY 5	Shared a TIN with another primary care practice	913,196 (85.6%)	$1.7 ($1.4)	1.2%		416,348 (88.3%)	-$1.8 ($2.1)	-1.2%		496,945 (83.5%)	$5.5*** ($1.9)	3.9%	
	Did not share a TIN with another primary care practice	153,630 (14.4%)	$1.0 ($2.9)	0.7%	0.29	55,289 (11.7%)	-$3.5 ($4.3)	-2.3%	0.24	98,244 (16.5%)	$3.4 ($3.7)	2.5%	0.49
Practice type: multi-specialty versus primary care only													
PY 1 through PY 5	Multi-specialty	278,801 (26.1%)	$2.8 ($2.8)	1.9%		116,601 (24.7%)	-$2.2 ($4.4)	-1.4%		162,149 (27.2%)	$7.3** ($3.6)	5.2%	
	Primary care only	788,025 (73.9%)	$1.1 ($1.5)	0.8%	0.51	355,036 (75.3%)	-$2.2 ($2.2)	-1.5%	0.75	433,040 (72.8%)	$4.5** ($2.0)	3.2%	0.48
Urbanicity of practice's county: rural or suburban versus urban													
PY 1 through PY 5	Rural	82,613 (7.7%)	$10.8** ($4.5)	8.5%		18,533 (3.9%)	$8.3 ($6.5)	6.5%		63,941 (10.7%)	$11.1** ($5.4)	8.7%	
	Suburban	170,323 (16.0%)	$5.7* ($3.1)	4.1%		75,938 (16.1%)	$2.9 ($4.7)	2.0%		94,390 (15.9%)	$9.5** ($4.3)	7.2%	
	Urban	813,890 (76.3%)	-$0.3 ($1.5)	-0.2%	0.03	377,166 (80.0%)	-$3.7* ($2.2)	-2.4%	0.16	436,858 (73.4%)	$3.4* ($2.0)	2.4%	0.22

Table 5.J.9. ***(continued)***

	Practice subgroup definition, based on baseline characteristics	Track 2– Overall				Track 2–SSP				Track 2–Non-SSP			
		Number (percentage) of CPC+ beneficiaries in subgroup at baseline	Impact estimate (standard error)	Percentage impact	*p*-Value for difference in impact estimates between subgroups[a]	Number (percentage) of CPC+ beneficiaries in subgroup at baseline	Impact estimate (standard error)	Percentage impact	*p*-Value for difference in impact estimates between subgroups[a]	Number (percentage) of CPC+ beneficiaries in subgroup at baseline	Impact estimate (standard error)	Percentage impact	*p*-Value for difference in impact estimates between subgroups[a]
Monthly Medicare expenditures for medical hospital admissions (per beneficiary per month)													
Whether practice participated in prior primary care transformation initiatives (recognized as a medical home or participated in MAPCP or CPC Classic)													
PY 1 through PY 5	Yes	865,798 (81.2%)	-$5.3*** ($1.3)	-3.6%		385,875 (81.8%)	-$5.8*** ($2.0)	-3.9%		479,947 (80.6%)	-$4.7*** ($1.7)	-3.3%	
	No	201,028 (18.8%)	$1.7 ($2.2)	1.1%	0.00	85,762 (18.2%)	$6.5 ($4.1)	3.9%	0.00	115,242 (19.4%)	-$2.5 ($2.4)	-1.7%	0.34
Large and medium, versus small practice based on number of primary care practitioners													
PY 1 through PY 5	Large (6+ primary care practitioners)	589,224 (55.2%)	-$4.4*** ($1.6)	-2.9%		279,067 (59.2%)	-$4.1* ($2.5)	-2.7%		310,301 (52.1%)	-$3.8* ($2.1)	-2.7%	
	Medium (3–5 primary care practitioners)	340,406 (31.9%)	-$2.7 ($1.9)	-1.9%		134,103 (28.4%)	-$0.9 ($3.1)	-0.7%		206,177 (34.6%)	-$4.9** ($2.3)	-3.4%	
	Small (1–2 primary care practitioners)	137,196 (12.9%)	-$5.2* ($2.9)	-3.3%	0.35	58,467 (12.4%)	-$7.0 ($4.9)	-4.3%	0.13	78,712 (13.2%)	-$4.3 ($3.5)	-2.8%	0.98
Whether hospital- or system-owned													
PY 1 through PY 5	Hospital-or system-owned	619,957 (58.1%)	-$2.0 ($1.5)	-1.4%		289,350 (61.4%)	-$1.9 ($2.3)	-1.3%		330,724 (55.6%)	-$2.3 ($2.0)	-1.6%	
	Independent	446,869 (41.9%)	-$6.6*** ($1.7)	-4.5%	0.02	182,287 (38.6%)	-$6.2** ($2.8)	-4.0%	0.07	264,465 (44.4%)	-$6.7*** ($2.0)	-4.7%	0.09
Whether the practice shared a TIN with another primary care practice													
PY 1 through PY 5	Shared a TIN with another primary care practice	913,196 (85.6%)	-$3.7*** ($1.3)	-2.5%		416,348 (88.3%)	-$3.3* ($1.9)	-2.2%		496,945 (83.5%)	-$4.0** ($1.6)	-2.7%	
	Did not share a TIN with another primary care practice	153,630 (14.4%)	-$4.3* ($2.5)	-3.0%	0.39	55,289 (11.7%)	-$4.5 ($4.9)	-3.0%	0.42	98,244 (16.5%)	-$4.4 ($2.8)	-3.2%	0.72
Practice type: multi-specialty versus primary care only													
PY 1 through PY 5	Multi-specialty	278,801 (26.1%)	-$7.5*** ($2.6)	-4.9%		116,601 (24.7%)	-$10.7** ($4.4)	-6.6%		162,149 (27.2%)	-$3.4 ($3.3)	-2.4%	
	Primary care only	788,025 (73.9%)	-$2.7** ($1.2)	-1.9%	0.07	355,036 (75.3%)	-$1.3 ($1.9)	-0.8%	0.01	433,040 (72.8%)	-$4.6*** ($1.6)	-3.2%	0.85
Urbanicity of practice's county: rural or suburban versus urban													
PY 1 through PY 5	Rural	82,613 (7.7%)	-$5.0 ($3.4)	-3.8%		18,533 (3.9%)	-$7.3 ($5.8)	-5.6%		63,941 (10.7%)	-$3.4 ($4.1)	-2.6%	
	Suburban	170,323 (16.0%)	-$3.9 ($3.3)	-2.8%		75,938 (16.1%)	-$4.1 ($5.0)	-2.9%		94,390 (15.9%)	-$1.9 ($4.4)	-1.4%	
	Urban	813,890 (76.3%)	-$3.9*** ($1.3)	-2.6%	0.99	377,166 (80.0%)	-$3.3 ($2.0)	-2.1%	0.85	436,858 (73.4%)	-$4.9*** ($1.6)	-3.3%	0.79

Table 5.J.9. ***(continued)***

		Track 2– Overall				Track 2–SSP				Track 2–Non-SSP			
	Practice subgroup definition, based on baseline characteristics	Number (percentage) of CPC+ beneficiaries in subgroup at baseline	Impact estimate (standard error)	Percentage impact	*p*-Value for difference in impact estimates between subgroups[a]	Number (percentage) of CPC+ beneficiaries in subgroup at baseline	Impact estimate (standard error)	Percentage impact	*p*-Value for difference in impact estimates between subgroups[a]	Number (percentage) of CPC+ beneficiaries in subgroup at baseline	Impact estimate (standard error)	Percentage impact	*p*-Value for difference in impact estimates between subgroups[a]
Annualized acute surgical hospital admissions (per 1,000 beneficiaries per year)													
Whether practice participated in prior primary care transformation initiatives (recognized as a medical home or participated in MAPCP or CPC Classic)													
PY 1 through PY 5	Yes	865,798 (81.2%)	-0.2 (0.7)	-0.2%		385,875 (81.8%)	-1.5* (0.9)	-1.9%		479,947 (80.6%)	1.2 (0.9)	1.5%	
	No	201,028 (18.8%)	0.6 (1.0)	0.7%	0.65	85,762 (18.2%)	2.4 (1.6)	3.1%	0.06	115,242 (19.4%)	-0.7 (1.3)	-0.9%	0.26
Large and medium, versus small practice based on number of primary care practitioners													
PY 1 through PY 5	Large (6+ primary care practitioners)	589,224 (55.2%)	-0.6 (0.8)	-0.7%		279,067 (59.2%)	-1.2 (1.1)	-1.5%		310,301 (52.1%)	0.2 (1.1)	0.3%	
	Medium (3–5 primary care practitioners)	340,406 (31.9%)	0.6 (1.0)	0.7%		134,103 (28.4%)	-0.9 (1.4)	-1.1%		206,177 (34.6%)	1.9 (1.5)	2.4%	
	Small (1–2 primary care practitioners)	137,196 (12.9%)	0.7 (1.4)	0.9%	0.34	58,467 (12.4%)	1.3 (2.1)	1.7%	0.78	78,712 (13.2%)	0.3 (1.9)	0.4%	0.28
Whether hospital- or system-owned													
PY 1 through PY 5	Hospital-or system-owned	619,957 (58.1%)	0.1 (0.7)	0.2%		289,350 (61.4%)	-1.1 (1.0)	-1.4%		330,724 (55.6%)	1.4 (1.1)	1.8%	
	Independent	446,869 (41.9%)	-0.3 (0.9)	-0.3%	0.58	182,287 (38.6%)	-0.3 (1.3)	-0.4%	0.70	264,465 (44.4%)	0.1 (1.2)	0.1%	0.31
Whether the practice shared a TIN with another primary care practice													
PY 1 through PY 5	Shared a TIN with another primary care practice	913,196 (85.6%)	0.1 (0.6)	0.1%		416,348 (88.3%)	-0.3 (0.9)	-0.3%		496,945 (83.5%)	0.7 (0.9)	0.8%	
	Did not share a TIN with another primary care practice	153,630 (14.4%)	-0.6 (1.3)	-0.8%	0.37	55,289 (11.7%)	-3.8** (1.9)	-4.6%	0.02	98,244 (16.5%)	0.9 (1.8)	1.2%	0.82
Practice type: multi-specialty versus primary care only													
PY 1 through PY 5	Multi-specialty	278,801 (26.1%)	1.3 (1.1)	1.6%		116,601 (24.7%)	-0.3 (1.6)	-0.4%		162,149 (27.2%)	2.8* (1.6)	3.6%	
	Primary care only	788,025 (73.9%)	-0.5 (0.7)	-0.6%	0.08	355,036 (75.3%)	-1.0 (1.0)	-1.2%	0.45	433,040 (72.8%)	0.1 (0.9)	0.1%	0.10
Urbanicity of practice's county: rural or suburban versus urban													
PY 1 through PY 5	Rural	82,613 (7.7%)	3.9** (1.8)	5.3%		18,533 (3.9%)	1.3 (3.3)	1.6%		63,941 (10.7%)	4.1* (2.2)	5.7%	
	Suburban	170,323 (16.0%)	0.4 (1.6)	0.5%		75,938 (16.1%)	-0.6 (2.4)	-0.7%		94,390 (15.9%)	1.8 (2.3)	2.3%	
	Urban	813,890 (76.3%)	-0.5 (0.6)	-0.6%	0.07	377,166 (80.0%)	-1.0 (0.9)	-1.2%	0.81	436,858 (73.4%)	0.1 (0.9)	0.2%	0.22

Table 5.J.9. ***(continued)***

	Practice subgroup definition, based on baseline characteristics	Track 2– Overall				Track 2–SSP				Track 2–Non-SSP			
		Number (percentage) of CPC+ beneficiaries in subgroup at baseline	Impact estimate (standard error)	Percentage impact	*p*-Value for difference in impact estimates between subgroups[a]	Number (percentage) of CPC+ beneficiaries in subgroup at baseline	Impact estimate (standard error)	Percentage impact	*p*-Value for difference in impact estimates between subgroups[a]	Number (percentage) of CPC+ beneficiaries in subgroup at baseline	Impact estimate (standard error)	Percentage impact	*p*-Value for difference in impact estimates between subgroups[a]
Annualized acute medical hospital admissions (per 1,000 beneficiaries per year)													
Whether practice participated in prior primary care transformation initiatives (recognized as a medical home or participated in MAPCP or CPC Classic)													
PY 1 through PY 5	Yes	865,798 (81.2%)	-3.9*** (1.5)	-2.0%		385,875 (81.8%)	-3.3 (2.2)	-1.7%		479,947 (80.6%)	-4.0** (2.0)	-2.1%	
	No	201,028 (18.8%)	1.1 (2.5)	0.6%	0.09	85,762 (18.2%)	9.7** (4.2)	4.6%	0.01	115,242 (19.4%)	-5.8** (2.8)	-3.1%	0.56
Large and medium, versus small practice based on number of primary care practitioners													
PY 1 through PY 5	Large (6+ primary care practitioners)	589,224 (55.2%)	-5.2*** (1.8)	-2.7%		279,067 (59.2%)	-3.0 (2.7)	-1.4%		310,301 (52.1%)	-6.2*** (2.3)	-3.3%	
	Medium (3–5 primary care practitioners)	340,406 (31.9%)	0.1 (2.2)	0.1%		134,103 (28.4%)	1.6 (3.2)	0.9%		206,177 (34.6%)	-1.4 (2.9)	-0.8%	
	Small (1–2 primary care practitioners)	137,196 (12.9%)	-1.1 (3.1)	-0.5%	0.38	58,467 (12.4%)	2.9 (4.8)	1.5%	0.97	78,712 (13.2%)	-4.4 (4.0)	-2.2%	0.21
Whether hospital- or system-owned													
PY 1 through PY 5	Hospital-or system-owned	619,957 (58.1%)	-0.1 (1.7)	0.0%		289,350 (61.4%)	1.4 (2.4)	0.7%		330,724 (55.6%)	-1.2 (2.3)	-0.6%	
	Independent	446,869 (41.9%)	-7.0*** (1.9)	-3.7%	0.00	182,287 (38.6%)	-4.6 (3.1)	-2.3%	0.04	264,465 (44.4%)	-8.2*** (2.4)	-4.6%	0.03
Whether the practice shared a TIN with another primary care practice													
PY 1 through PY 5	Shared a TIN with another primary care practice	913,196 (85.6%)	-2.4* (1.4)	-1.2%		416,348 (88.3%)	-0.5 (2.1)	-0.2%		496,945 (83.5%)	-3.5* (1.9)	-1.9%	
	Did not share a TIN with another primary care practice	153,630 (14.4%)	-4.3 (2.7)	-2.4%	0.31	55,289 (11.7%)	-2.6 (5.1)	-1.3%	0.32	98,244 (16.5%)	-5.8* (3.1)	-3.3%	0.68
Practice type: multi-specialty versus primary care only													
PY 1 through PY 5	Multi-specialty	278,801 (26.1%)	-5.7* (3.0)	-2.9%		116,601 (24.7%)	-8.8* (4.7)	-4.1%		162,149 (27.2%)	-1.9 (4.1)	-1.0%	
	Primary care only	788,025 (73.9%)	-2.0 (1.4)	-1.1%	0.62	355,036 (75.3%)	1.6 (2.1)	0.8%	0.03	433,040 (72.8%)	-5.2*** (1.8)	-2.7%	0.24
Urbanicity of practice's county: rural or suburban versus urban													
PY 1 through PY 5	Rural	82,613 (7.7%)	-7.2* (4.0)	-4.2%		18,533 (3.9%)	-7.4 (6.7)	-4.1%		63,941 (10.7%)	-7.0 (4.9)	-4.2%	
	Suburban	170,323 (16.0%)	-6.0 (3.7)	-3.1%		75,938 (16.1%)	-8.6 (5.4)	-4.4%		94,390 (15.9%)	-1.2 (5.4)	-0.7%	
	Urban	813,890 (76.3%)	-1.9 (1.4)	-1.0%	0.45	377,166 (80.0%)	0.9 (2.2)	0.5%	0.14	436,858 (73.4%)	-4.6** (1.9)	-2.4%	0.83

Source: Mathematica's analysis of Medicare claims data from January 2016 through December 2021.

Note: The estimates (and standard errors) in the impact estimate column show the separate subgroup-specific impacts over the first four years of CPC+ for each practice characteristic listed in the table.

Table 5.J.9. ***(continued)***

[a] The p-values in the last column represent the results from tests for statistically significant differences in impact estimates between the subgroups based on the baseline practice characteristic (using a t-test for subgroups with two categories and an F-test for subgroups with more than two categories).

*/**/***Within-subgroup estimate significantly different from zero at the 0.10/0.05/0.01 level, two-tailed test.

CPC = Comprehensive Primary Care; MAPCP = Multi-Payer Advanced Primary Care Practice; PY = Program Year; SSP = Medicare Shared Savings Program; TIN = Tax Identification Number.

E. Impact on hospitalizations by beneficiary characteristics

Across the five program years, CPC+ reduced expenditures for medical hospital admissions more for beneficiaries in the highest quartile of HCC score distribution compared with remaining beneficiaries in both tracks. In Track 1, CPC+ reduced expenditures for medical admissions by $7.9 (2.1 percent of baseline expenditures) for beneficiaries in the highest quartile of the HCC score distribution versus $2.5 (2.8 percent of baseline values) for beneficiaries who were not in the highest quartile. This difference was statistically significant ($p = 0.07$; Table 5.J.10). There were also statistical differences in Track 2: CPC+ reduced expenditures for medical admissions by $8.9 (2.3 percent) for beneficiaries in the highest quartile of the HCC score distribution versus $2.0 (2.3 percent) for remaining beneficiaries ($p = 0.04$ for the difference in impact estimates). SSP practices drove the difference in Track 2, where CPC+ reduced expenditures for medical admissions only for beneficiaries in the highest quartile of the HCC score distribution ($10.4), with no impacts for beneficiaries not in the highest quartile (Table 5.J.11). When grouping beneficiaries based on being in the highest quartile of the HCC score distribution or with dementia, we see a similar pattern using this definition—where greater reductions were concentrated in the high-risk group—but the differences were significant only in Track 2 (Table 5.J.11).

In Track 2, there were differences in impact estimates based on dual eligibility for Medicare and Medicaid, with impacts concentrated among patients who were dually eligible. Over the five years of the program, CPC+ reduced expenditures for medical hospitalizations by $10.2 PBPM (3.7 percent) for beneficiaries dually eligible for Medicare and Medicaid, versus $2.6 PBPM (1.9 percent) for beneficiaries who are not dually eligible (Table 5.J.11). This difference in impact estimates was statistically significant ($p = 0.054$). These differences were concentrated among SSP practices, where there were reductions of $15.1 PBPM (5.3 percent) for patients who are dually eligible, with no impact for beneficiaries who are not dually eligible. We observe a similar pattern of results for medical hospitalizations, with large reductions among duals of 12.6 admissions per 1,000 beneficiaries per year (3.6 percent) but no impact for those who are not dually eligible. The difference in impact estimates was statistically significant ($p < 0.01$). Differences were concentrated among SSP practices, where there were reductions of 17.3 admissions per 1,000 beneficiaries per year (4.8 percent) for patients who are dually eligible, with no impact for those who are not dually eligible. In addition, CPC+ increased the number of surgical admissions for dually eligible patients by 2.6 hospitalizations per 1,000 beneficiaries per year (3.0 percent). This result was statistically significant, but CPC+ had no impact for those who are not dually eligible.

For the other outcomes and other subgroups not discussed in this section, there were no statistical differences in the impact estimates between beneficiary subgroups (that is, the p-values were larger than 0.1).

Table 5.J.10. Estimates of five-year impacts of CPC+ on surgical and medical hospitalizations for Track 1, by baseline beneficiary characteristics

		Track 1–Overall				Track 1–SSP				Track 1– Non-SSP			
	Beneficiary subgroup definition, based on baseline characteristics	Number (percentage) of CPC+ beneficiaries in subgroup at baseline	Impact estimate (standard error)	Percentage impact	*p*-Value for difference in impact estimates between subgroups[a]	Number (percentage) of CPC+ beneficiaries in subgroup at baseline	Impact estimate (standard error)	Percentage impact	*p*-Value for difference in impact estimates between subgroups[a]	Number (percentage) of CPC+ beneficiaries in subgroup at baseline	Impact estimate (standard error)	Percentage impact	*p*-Value for difference in impact estimates between subgroups[a]
Monthly Medicare expenditures for surgical hospital admissions (per beneficiary per month)													
Patients in the highest quartile of the HCC score distribution													
PY 1 through 5	Yes	203,811 (25.9)	$0.9 ($3.8)	0.3%		115,215 (26.8)	-$2.7 ($5.0)	-1.0%		88,864 (25.0)	$4.9 ($5.9)	1.8%	
	No	583,156 (74.1)	$0.2 ($1.1)	0.2%	0.86	315,425 (73.2)	-$0.8 ($1.5)	-0.7%	0.72	266,666 (75.0)	$1.2 ($1.5)	1.1%	0.54
Patients in the highest decile of the HCC score distribution or who have dementia													
PY 1 through 5	Yes	123,085 (15.6)	-$1.1 ($5.0)	-0.4%		68,759 (16.0)	-$3.3 ($6.7)	-1.1%		54,382 (15.3)	$1.5 ($7.7)	0.5%	
	No	663,882 (84.4)	$0.7 ($1.1)	0.5%	0.73	361,881 (84.0)	-$0.9 ($1.6)	-0.7%	0.72	301,148 (84.7)	$2.2 ($1.6)	1.8%	0.92
Patients with anxiety or depression or substance use disorders													
PY 1 through 5	Yes	120,562 (16.6)	-$0.5 ($3.7)	-0.3%		66,746 (16.8)	-$0.9 ($5.1)	-0.5%		53,792 (16.4)	-$0.2 ($5.3)	-0.1%	
	No	604,012 (83.4)	$1.4 ($1.3)	1.0%	0.60	329,703 (83.2)	-$0.3 ($1.9)	-0.2%	0.91	273,568 (83.6)	$3.3* ($1.9)	2.4%	0.52
Patients with multiple chronic conditions (at least 2 of 12 frequently occurring chronic conditions[b]) and one or more hospitalizations[c]													
PY 1 through 5	Yes	68,204 (8.7)	$8.5 ($7.4)	2.5%		38,153 (8.9)	$9.4 ($10.1)	2.7%		30,089 (8.5)	$7.4 ($11.0)	2.2%	
	No	718,763 (91.3)	-$0.4 ($1.2)	-0.3%	0.23	392,487 (91.1)	-$2.3 ($1.6)	-1.7%	0.24	325,442 (91.5)	$1.6 ($1.7)	1.2%	0.60
Patients dually eligible for Medicare and Medicaid													
PY 1 through 5	Yes	107,885 (12.6)	$0.3 ($4.0)	0.2%		55,728 (11.9)	-$2.6 ($5.6)	-1.5%		51,626 (13.3)	$3.0 ($5.6)	1.8%	
	No	746,776 (87.4)	$0.1 ($1.3)	0.1%	0.97	410,653 (88.1)	-$2.0 ($1.8)	-1.3%	0.92	335,619 (86.7)	$2.4 ($1.9)	1.7%	0.91
Monthly Medicare expenditures for medical hospital admissions (per beneficiary per month)													
Patients in the highest quartile of the HCC score distribution													
PY 1 through 5	Yes	203,811 (25.9)	-$7.9** ($3.1)	-2.1%		115,215 (26.8)	-$9.9** ($4.1)	-2.6%		88,864 (25.0)	-$5.7 ($4.7)	-1.5%	
	No	583,156 (74.1)	-$2.5*** ($0.8)	-2.8%	0.07	315,425 (73.2)	-$3.5*** ($1.0)	-3.9%	0.12	266,666 (75.0)	-$1.5 ($1.1)	-1.7%	0.37

Table 5.J.10. ***(continued)***

		Track 1–Overall				Track 1–SSP				Track 1– Non-SSP			
	Beneficiary subgroup definition, based on baseline characteristics	Number (percentage) of CPC+ beneficiaries in subgroup at baseline	Impact estimate (standard error)	Percentage impact	*p*-Value for difference in impact estimates between subgroups[a]	Number (percentage) of CPC+ beneficiaries in subgroup at baseline	Impact estimate (standard error)	Percentage impact	*p*-Value for difference in impact estimates between subgroups[a]	Number (percentage) of CPC+ beneficiaries in subgroup at baseline	Impact estimate (standard error)	Percentage impact	*p*-Value for difference in impact estimates between subgroups[a]
Patients in the highest decile of the HCC score distribution or who have dementia													
PY 1 through 5	Yes	123,085 (15.6)	-$8.5* ($4.5)	-1.8%		68,759 (16.0)	-$6.4 ($5.9)	-1.4%		54,382 (15.3)	-$10.6 ($6.7)	-2.3%	
	No	663,882 (84.4)	-$3.3*** ($0.8)	-3.0%	0.24	361,881 (84.0)	-$5.1*** ($1.1)	-4.6%	0.83	301,148 (84.7)	-$1.4 ($1.2)	-1.4%	0.17
Patients with anxiety or depression or substance use disorders													
PY 1 through 5	Yes	120,562 (16.6)	-$0.2 ($3.0)	-0.1%		66,746 (16.8)	-$4.1 ($3.9)	-1.8%		53,792 (16.4)	$4.0 ($4.7)	1.9%	
	No	604,012 (83.4)	-$3.7*** ($1.0)	-3.0%	0.25	329,703 (83.2)	-$4.5*** ($1.3)	-3.5%	0.91	273,568 (83.6)	-$2.9** ($1.4)	-2.4%	0.15
Patients with multiple chronic conditions (at least 2 of 12 frequently occurring chronic conditions[b]) and one or more hospitalizations[c]													
PY 1 through 5	Yes	68,204 (8.7)	-$8.1 ($6.5)	-1.5%		38,153 (8.9)	-$4.1 ($8.7)	-0.8%		30,089 (8.5)	-$12.4 ($9.8)	-2.3%	
	No	718,763 (91.3)	-$3.5*** ($0.9)	-2.8%	0.47	392,487 (91.1)	-$5.3*** ($1.2)	-4.1%	0.89	325,442 (91.5)	-$1.6 ($1.2)	-1.4%	0.26
Patients dually eligible for Medicare and Medicaid													
PY 1 through 5	Yes	107,885 (12.6)	-$5.0 ($3.4)	-1.9%		55,728 (11.9)	-$10.9** ($4.9)	-4.1%		51,626 (13.3)	$0.9 ($4.7)	0.4%	
	No	746,776 (87.4)	-$2.7*** ($0.9)	-2.1%	0.50	410,653 (88.1)	-$3.6*** ($1.2)	-2.7%	0.13	335,619 (86.7)	-$1.9 ($1.4)	-1.6%	0.54
Annualized acute surgical hospital admissions (per 1,000 beneficiaries per year)													
Patients in the highest quartile of the HCC score distribution													
PY 1 through 5	Yes	203,811 (25.9)	1.4 (1.5)	1.0%		115,215 (26.8)	0.3 (1.8)	0.2%		88,864 (25.0)	2.7 (2.5)	1.9%	
	No	583,156 (74.1)	-0.1 (0.5)	-0.1%	0.34	315,425 (73.2)	-0.5 (0.7)	-0.8%	0.66	266,666 (75.0)	0.4 (0.8)	0.6%	0.37
Patients in the highest decile of the HCC score distribution or who have dementia													
PY 1 through 5	Yes	123,085 (15.6)	0.0 (1.9)	0.0%		68,759 (16.0)	-1.0 (2.4)	-0.7%		54,382 (15.3)	1.1 (3.0)	0.7%	
	No	663,882 (84.4)	0.4 (0.5)	0.5%	0.83	361,881 (84.0)	-0.2 (0.7)	-0.2%	0.74	301,148 (84.7)	1.0 (0.8)	1.3%	0.97
Patients with anxiety or depression or substance use disorders													
PY 1 through 5	Yes	120,562 (16.6)	-1.6 (1.6)	-1.5%		66,746 (16.8)	-1.5 (2.2)	-1.4%		53,792 (16.4)	-1.7 (2.3)	-1.6%	
	No	604,012 (83.4)	1.1* (0.6)	1.4%	0.11	329,703 (83.2)	0.3 (0.8)	0.4%	0.42	273,568 (83.6)	1.8** (0.9)	2.4%	0.15

Table 5.J.10. ***(continued)***

		Track 1–Overall				Track 1–SSP				Track 1– Non-SSP			
	Beneficiary subgroup definition, based on baseline characteristics	Number (percentage) of CPC+ beneficiaries in subgroup at baseline	Impact estimate (standard error)	Percentage impact	*p*-Value for difference in impact estimates between subgroups[a]	Number (percentage) of CPC+ beneficiaries in subgroup at baseline	Impact estimate (standard error)	Percentage impact	*p*-Value for difference in impact estimates between subgroups[a]	Number (percentage) of CPC+ beneficiaries in subgroup at baseline	Impact estimate (standard error)	Percentage impact	*p*-Value for difference in impact estimates between subgroups[a]
Patients with multiple chronic conditions (at least 2 of 12 frequently occurring chronic conditions[b]) and one or more hospitalizations[c]													
PY 1 through 5	Yes	68,204 (8.7)	2.8 (2.9)	1.7%		38,153 (8.9)	1.3 (3.6)	0.8%		30,089 (8.5)	4.5 (4.6)	2.7%	
	No	718,763 (91.3)	0.0 (0.5)	0.1%	0.34	392,487 (91.1)	-0.5 (0.7)	-0.7%	0.63	325,442 (91.5)	0.6 (0.8)	0.8%	0.39
Patients dually eligible for Medicare and Medicaid													
PY 1 through 5	Yes	107,885 (12.6)	0.7 (1.6)	0.8%		55,728 (11.9)	-1.4 (2.1)	-1.6%		51,626 (13.3)	2.9 (2.4)	3.3%	
	No	746,776 (87.4)	-0.2 (0.6)	-0.2%	0.60	410,653 (88.1)	-0.7 (0.7)	-0.9%	0.76	335,619 (86.7)	0.4 (0.9)	0.6%	0.35
Annualized acute medical hospital admissions (per 1,000 beneficiaries per year)													
Patients in the highest quartile of the HCC score distribution													
PY 1 through 5	Yes	203,811 (25.9)	-4.4 (3.8)	-0.9%		115,215 (26.8)	-5.0 (4.7)	-1.0%		88,864 (25.0)	-3.2 (6.0)	-0.6%	
	No	583,156 (74.1)	-1.9** (0.9)	-1.6%	0.50	315,425 (73.2)	-3.1** (1.2)	-2.5%	0.67	266,666 (75.0)	-0.4 (1.3)	-0.4%	0.64
Patients in the highest decile of the HCC score distribution or who have dementia													
PY 1 through 5	Yes	123,085 (15.6)	-6.5 (5.4)	-1.1%		68,759 (16.0)	-3.1 (6.8)	-0.5%		54,382 (15.3)	-9.6 (8.5)	-1.6%	
	No	663,882 (84.4)	-2.0** (1.0)	-1.4%	0.41	361,881 (84.0)	-3.8*** (1.4)	-2.6%	0.92	301,148 (84.7)	0.1 (1.5)	0.1%	0.25
Patients with anxiety or depression or substance use disorders													
PY 1 through 5	Yes	120,562 (16.6)	0.3 (3.8)	0.1%		66,746 (16.8)	-2.6 (4.9)	-0.9%		53,792 (16.4)	3.9 (5.9)	1.3%	
	No	604,012 (83.4)	-2.1* (1.1)	-1.3%	0.53	329,703 (83.2)	-2.8* (1.5)	-1.7%	0.97	273,568 (83.6)	-1.1 (1.7)	-0.7%	0.39
Patients with multiple chronic conditions (at least 2 of 12 frequently occurring chronic conditions[b]) and one or more hospitalizations[c]													
PY 1 through 5	Yes	68,204 (8.7)	-5.1 (7.9)	-0.7%		38,153 (8.9)	-1.0 (10.1)	-0.1%		30,089 (8.5)	-9.0 (12.3)	-1.3%	
	No	718,763 (91.3)	-2.4** (1.1)	-1.5%	0.73	392,487 (91.1)	-4.0*** (1.4)	-2.4%	0.76	325,442 (91.5)	-0.5 (1.6)	-0.3%	0.48
Patients dually eligible for Medicare and Medicaid													
PY 1 through 5	Yes	107,885 (12.6)	-5.0 (4.0)	-1.5%		55,728 (11.9)	-13.7** (5.6)	-4.0%		51,626 (13.3)	4.0 (5.6)	1.2%	
	No	746,776 (87.4)	-2.0* (1.2)	-1.2%	0.45	410,653 (88.1)	-2.3 (1.5)	-1.3%	0.04	335,619 (86.7)	-1.4 (1.8)	-0.9%	0.33

Source: Mathematica's analysis of Medicare claims data from January 2016 through December 2021.

Table 5.J.10. ***(continued)***

Note: To determine subgroup membership, we measured beneficiary characteristics at the start of the year-long baseline period for baseline observations and at the start of PY 1 for observations in the intervention period (PY 1 through PY 5). The estimates (and standard errors) in the impact estimate column show the separate subgroup-specific impacts for each beneficiary characteristic listed in the table. We could not observe diagnoses (which are used to determine HCCs and to calculate HCC scores) at baseline for beneficiaries who were new to Medicare during the program years; we therefore excluded new Medicare beneficiaries from all subgroup analyses (except the analysis based on dual status, because beneficiaries who are new to Medicare cannot, by definition, be enrolled in both Medicare and Medicaid before joining Medicare). This process excluded about 10 percent of the observations from the regressions for the subgroups defined by HCC scores and chronic conditions.

[a] The p-values represent results from tests for statistically significant differences in impact estimates between the subgroups based on the baseline beneficiary characteristic (using a t-test for all subgroups).

[b] The 12 frequently occurring chronic conditions are (1) congestive heart failure, (2) chronic obstructive pulmonary disease, (3) history of acute myocardial infarction, (4) ischemic heart disease, (5) diabetes, (6) metastatic cancer or acute leukemia, (7) history of stroke, (8) depression, (9) dementia, (10) atrial fibrillation, (11) rheumatoid arthritis or osteoarthritis, and (12) chronic kidney disease.

[c] For observations in the baseline year, we measured hospitalizations in 2015, the year before the start of the baseline year. For observations in the intervention period, we measured hospitalizations in 2016, the year before the start of PY 1.

*/**/*** Significantly different from zero at the 0.10/0.05/0.01 level, two-tailed test.

HCC = hierarchical condition category; PY = Program Year.

Table 5.J.11. Estimates of five-year impacts of CPC+ on surgical and medical hospitalizations for Track 2, by baseline beneficiary characteristics

		Track 2–Overall				Track 2–SSP				Track 2–Non-SSP			
	Beneficiary subgroup definition, based on baseline characteristics	Number (percentage) of CPC+ beneficiaries in subgroup at baseline	Impact estimate (standard error)	Percentage impact	*p*-Value for difference in impact estimate between subgroups[a]	Number (percentage) of CPC+ beneficiaries in subgroup at baseline	Impact estimate (standard error)	Percentage impact	*p*-Value for difference in impact estimate between subgroups[a]	Number (percentage) of CPC+ beneficiaries in subgroup at baseline	Impact estimate (standard error)	Percentage impact	*p*-Value for difference in impact estimate between subgroups[a]
Monthly Medicare expenditures for surgical hospital admissions (per beneficiary per month)													
Patients in the highest quartile of the HCC score distribution													
PY 1 through 5	Yes	268,430 (26.1)	$7.0* ($3.8)	2.6%		120,947 (26.8)	-$3.7 ($5.5)	-1.3%		146,522 (25.5)	$15.9*** ($5.1)	6.0%	
	No	761,970 (73.9)	$1.4 ($1.1)	1.2%	0.15	330,277 (73.2)	$0.0 ($1.7)	0.0%	0.51	428,947 (74.5)	$2.7* ($1.5)	2.5%	0.01
Patients in the highest decile of the HCC score distribution or who have dementia													
PY 1 through 5	Yes	162,510 (15.8)	$6.0 ($5.2)	2.0%		71,030 (15.7)	-$10.7 ($7.6)	-3.3%		90,895 (15.8)	$19.3*** ($7.1)	6.9%	
	No	867,891 (84.2)	$2.2* ($1.2)	1.7%	0.46	380,194 (84.3)	$0.7 ($1.8)	0.5%	0.14	484,574 (84.2)	$3.7** ($1.5)	2.9%	0.03
Patients with anxiety/depression or substance use disorders													
PY 1 through 5	Yes	164,048 (17.3)	-$0.1 ($3.8)	0.0%		74,382 (17.8)	-$7.8 ($5.8)	-3.9%		89,058 (16.8)	$6.2 ($5.0)	3.4%	
	No	784,877 (82.7)	$3.4** ($1.4)	2.5%	0.37	342,453 (82.2)	$0.9 ($2.1)	0.6%	0.13	439,501 (83.2)	$5.8*** ($1.8)	4.3%	0.93
Patients with multiple chronic conditions (at least 2 of 12 frequently occurring chronic conditions[b]) and one or more hospitalizations[c]													
PY 1 through 5	Yes	90,543 (8.8)	$10.0 ($7.7)	2.9%		41,080 (9.1)	$0.7 ($11.9)	0.2%		49,139 (8.5)	$17.9* ($10.2)	5.4%	
	No	939,858 (91.2)	$2.2* ($1.2)	1.6%	0.31	410,144 (90.9)	-$1.1 ($1.8)	-0.8%	0.88	526,331 (91.5)	$4.9*** ($1.6)	3.8%	0.20
Patients dually eligible for Medicare and Medicaid													
PY 1 through 5	Yes	140,782 (12.5)	$6.5 ($3.9)	3.7%		55,837 (11.3)	$7.3 ($6.3)	4.1%		84,414 (13.5)	$6.3 ($5.0)	3.7%	
	No	984,688 (87.5)	$1.1 ($1.4)	0.8%	0.20	438,154 (88.7)	-$3.4 ($2.1)	-2.3%	0.11	542,895 (86.5)	$5.1*** ($1.8)	3.7%	0.82
Monthly Medicare expenditures for medical hospital admissions (per beneficiary per month)													
Patients in the highest quartile of the HCC score distribution													
PY 1 through 5	Yes	268,430 (26.1)	-$8.9** ($3.5)	-2.3%		120,947 (26.8)	-$10.4* ($5.6)	-2.6%		146,522 (25.5)	-$7.9* ($4.3)	-2.1%	
	No	761,970 (73.9)	-$2.0** ($0.9)	-2.3%	0.04	330,277 (73.2)	-$0.5 ($1.4)	-0.5%	0.07	428,947 (74.5)	-$3.2*** ($1.1)	-3.7%	0.28

Table 5.J.11. ***(continued)***

		Track 2–Overall				Track 2–SSP				Track 2–Non-SSP			
	Beneficiary subgroup definition, based on baseline characteristics	Number (percentage) of CPC+ beneficiaries in subgroup at baseline	Impact estimate (standard error)	Percentage impact	*p*-Value for difference in impact estimate between subgroups[a]	Number (percentage) of CPC+ beneficiaries in subgroup at baseline	Impact estimate (standard error)	Percentage impact	*p*-Value for difference in impact estimate between subgroups[a]	Number (percentage) of CPC+ beneficiaries in subgroup at baseline	Impact estimate (standard error)	Percentage impact	*p*-Value for difference in impact estimate between subgroups[a]
Patients in the highest decile of the HCC score distribution or who have dementia													
PY 1 through 5	Yes	162,510 (15.8)	-$11.0** ($5.1)	-2.3%		71,030 (15.7)	-$17.7** ($8.4)	-3.5%		90,895 (15.8)	-$6.0 ($6.2)	-1.3%	
	No	867,891 (84.2)	-$2.6*** ($0.9)	-2.4%	0.09	380,194 (84.3)	-$0.5 ($1.5)	-0.5%	0.04	484,574 (84.2)	-$4.3*** ($1.1)	-4.0%	0.78
Patients with anxiety/depression or substance use disorders													
PY 1 through 5	Yes	164,048 (17.3)	-$4.4 ($3.3)	-2.0%		74,382 (17.8)	-$3.7 ($4.9)	-1.6%		89,058 (16.8)	-$5.2 ($4.3)	-2.3%	
	No	784,877 (82.7)	-$3.3*** ($1.1)	-2.6%	0.73	342,453 (82.2)	-$1.7 ($1.8)	-1.3%	0.68	439,501 (83.2)	-$4.5*** ($1.3)	-3.6%	0.88
Patients with multiple chronic conditions (at least 2 of 12 frequently occurring chronic conditions[b]) and one or more hospitalizations[c]													
PY 1 through 5	Yes	90,543 (8.8)	-$3.3 ($7.0)	-0.6%		41,080 (9.1)	-$3.2 ($10.6)	-0.6%		49,139 (8.5)	-$3.5 ($9.3)	-0.7%	
	No	939,858 (91.2)	-$3.8*** ($1.1)	-3.0%	0.94	410,144 (90.9)	-$2.9 ($1.8)	-2.2%	0.98	526,331 (91.5)	-$4.5*** ($1.2)	-3.6%	0.92
Patients dually eligible for Medicare and Medicaid													
PY 1 through 5	Yes	140,782 (12.5)	-$10.2** ($4.0)	-3.7%		55,837 (11.3)	-$15.1** ($6.6)	-5.3%		84,414 (13.5)	-$6.9 ($5.0)	-2.6%	
	No	984,688 (87.5)	-$2.6** ($1.1)	-1.9%	0.05	438,154 (88.7)	-$1.9 ($1.8)	-1.3%	0.04	542,895 (86.5)	-$3.2** ($1.3)	-2.5%	0.47
Annualized acute surgical hospital admissions (per 1,000 beneficiaries per year)													
Patients in the highest quartile of the HCC score distribution													
PY 1 through 5	Yes	268,430 (26.1)	0.5 (1.5)	0.4%		120,947 (26.8)	-0.9 (2.1)	-0.7%		146,522 (25.5)	1.8 (2.1)	1.3%	
	No	761,970 (73.9)	0.5 (0.5)	0.8%	0.99	330,277 (73.2)	0.1 (0.8)	0.2%	0.64	428,947 (74.5)	0.9 (0.7)	1.4%	0.68
Patients in the highest decile of the HCC score distribution or who have dementia													
PY 1 through 5	Yes	162,510 (15.8)	1.2 (2.0)	0.8%		71,030 (15.7)	-1.3 (2.9)	-0.9%		90,895 (15.8)	3.2 (2.7)	2.3%	
	No	867,891 (84.2)	0.4 (0.6)	0.6%	0.68	380,194 (84.3)	0.0 (0.8)	0.0%	0.65	484,574 (84.2)	0.8 (0.7)	1.1%	0.36
Patients with anxiety/depression or substance use disorders													
PY 1 through 5	Yes	164,048 (17.3)	0.0 (1.6)	0.0%		74,382 (17.8)	-0.7 (2.4)	-0.7%		89,058 (16.8)	0.6 (2.2)	0.6%	
	No	784,877 (82.7)	0.7 (0.6)	0.9%	0.68	342,453 (82.2)	0.0 (0.9)	0.0%	0.76	439,501 (83.2)	1.3 (0.8)	1.7%	0.77

Table 5.J.11. ***(continued)***

	Beneficiary subgroup definition, based on baseline characteristics	Track 2–Overall: Number (percentage) of CPC+ beneficiaries in subgroup at baseline	Track 2–Overall: Impact estimate (standard error)	Track 2–Overall: Percentage impact	Track 2–Overall: *p*-Value for difference in impact estimate between subgroups[a]	Track 2–SSP: Number (percentage) of CPC+ beneficiaries in subgroup at baseline	Track 2–SSP: Impact estimate (standard error)	Track 2–SSP: Percentage impact	Track 2–SSP: *p*-Value for difference in impact estimate between subgroups[a]	Track 2–Non-SSP: Number (percentage) of CPC+ beneficiaries in subgroup at baseline	Track 2–Non-SSP: Impact estimate (standard error)	Track 2–Non-SSP: Percentage impact	Track 2–Non-SSP: *p*-Value for difference in impact estimate between subgroups[a]
Patients with multiple chronic conditions (at least 2 of 12 frequently occurring chronic conditions[b]) and one or more hospitalizations[c]													
PY 1 through 5	Yes	90,543 (8.8)	2.0 (2.8)	1.2%		41,080 (9.1)	1.7 (4.0)	1.0%		49,139 (8.5)	2.4 (4.0)	1.4%	
	No	939,858 (91.2)	0.4 (0.6)	0.5%	0.57	410,144 (90.9)	-0.3 (0.8)	-0.4%	0.61	526,331 (91.5)	1.0 (0.8)	1.3%	0.73
Patients dually eligible for Medicare and Medicaid													
PY 1 through 5	Yes	140,782 (12.5)	2.6* (1.6)	3.0%		55,837 (11.3)	2.0 (2.4)	2.4%		84,414 (13.5)	3.0 (2.1)	3.5%	
	No	984,688 (87.5)	-0.3 (0.6)	-0.4%	0.08	438,154 (88.7)	-1.1 (0.9)	-1.4%	0.21	542,895 (86.5)	0.4 (0.8)	0.5%	0.24
Annualized acute medical hospital admissions (per 1,000 beneficiaries per year)													
Patients in the highest quartile of the HCC score distribution													
PY 1 through 5	Yes	268,430 (26.1)	-3.7 (3.9)	-0.7%		120,947 (26.8)	0.3 (5.9)	0.1%		146,522 (25.5)	-7.1 (5.1)	-1.5%	
	No	761,970 (73.9)	-1.5 (1.0)	-1.2%	0.57	330,277 (73.2)	1.1 (1.6)	0.9%	0.89	428,947 (74.5)	-3.4*** (1.2)	-2.9%	0.47
Patients in the highest decile of the HCC score distribution or who have dementia													
PY 1 through 5	Yes	162,510 (15.8)	-5.1 (5.6)	-0.8%		71,030 (15.7)	-4.9 (8.7)	-0.8%		90,895 (15.8)	-5.5 (7.3)	-0.9%	
	No	867,891 (84.2)	-1.6 (1.1)	-1.1%	0.52	380,194 (84.3)	2.0 (1.7)	1.3%	0.42	484,574 (84.2)	-4.3*** (1.3)	-3.0%	0.87
Patients with anxiety/depression or substance use disorders													
PY 1 through 5	Yes	164,048 (17.3)	-5.2 (3.9)	-1.7%		74,382 (17.8)	-3.4 (5.7)	-1.1%		89,058 (16.8)	-6.9 (5.4)	-2.3%	
	No	784,877 (82.7)	-1.1 (1.2)	-0.7%	0.30	342,453 (82.2)	2.7 (1.8)	1.6%	0.28	439,501 (83.2)	-4.1*** (1.5)	-2.5%	0.62
Patients with multiple chronic conditions (at least 2 of 12 frequently occurring chronic conditions[b]) and one or more hospitalizations[c]													
PY 1 through 5	Yes	90,543 (8.8)	-0.3 (8.4)	0.0%		41,080 (9.1)	-1.1 (12.6)	-0.1%		49,139 (8.5)	0.2 (11.3)	0.0%	
	No	939,858 (91.2)	-2.2* (1.1)	-1.3%	0.82	410,144 (90.9)	1.1 (1.8)	0.6%	0.86	526,331 (91.5)	-4.7*** (1.4)	-2.9%	0.66
Patients dually eligible for Medicare and Medicaid													
PY 1 through 5	Yes	140,782 (12.5)	-12.6*** (4.3)	-3.6%		55,837 (11.3)	-17.3** (6.8)	-4.8%		84,414 (13.5)	-9.3* (5.5)	-2.7%	
	No	984,688 (87.5)	-1.2 (1.2)	-0.7%	0.01	438,154 (88.7)	1.4 (1.9)	0.8%	0.01	542,895 (86.5)	-3.3** (1.6)	-1.9%	0.28

Source: Mathematica's analysis of Medicare claims data from January 2016 through December 2021.

Table 5.J.11. ***(continued)***

Note: To determine subgroup membership, we measured beneficiary characteristics at the start of the year-long baseline period for baseline observations and at the start of PY 1 for observations in the intervention period (PY 1 through PY 5). The estimates (and standard errors) in the impact estimate column show the separate subgroup-specific impacts for each beneficiary characteristic listed in the table. We could not observe diagnoses (which are used to determine HCCs and to calculate HCC scores) at baseline for beneficiaries who were new to Medicare during the program years; we therefore excluded new Medicare beneficiaries from all subgroup analyses (except the analysis based on dual status because beneficiaries who are new to Medicare cannot, by definition, be enrolled in both Medicare and Medicaid prior to joining Medicare). This process excludes about 10 percent of the observations from the regressions for the subgroups defined by HCC scores and chronic conditions.

[a] The *p*-values represent results from tests for statistically significant differences in impact estimates between the subgroups based on the baseline beneficiary characteristic (using a t-test for all subgroups).

[b] The 12 frequently occurring chronic conditions are (1) congestive heart failure, (2) chronic obstructive pulmonary disease, (3) history of acute myocardial infarction, (4) ischemic heart disease, (5) diabetes, (6) metastatic cancer or acute leukemia, (7) history of stroke, (8) depression, (9) dementia, (10) atrial fibrillation, (11) rheumatoid arthritis or osteoarthritis, and (12) chronic kidney disease.

[c] For observations in the baseline year, we measured hospitalizations in 2015, the year before the start of the baseline year. For observations in the intervention period, we measured hospitalizations in 2016, the year before the start of PY 1.

*/**/*** Significantly different from zero at the 0.10/0.05/0.01 level, two-tailed test.

HCC = hierarchical condition category; PY = Program Year.

5.J.4. Discussion

We report two key findings from this research examining the impact of CPC+ on types of acute hospitalizations. First, medical admissions drove the reductions in all-cause acute hospitalization rates and expenditures among CPC+ practices almost entirely. There were estimated impacts in multiple program years, with the largest effects in the third and fourth program years. Second, most of the reductions in medical hospitalizations came from admissions without an MCC or CC—that is, the least complex hospitalizations.

Our findings are generally consistent with CPC+ leading to improvements in timely access and enhanced care management activities. Prior research found that CPC+ practices in both tracks improved after-hours access since CPC+ began. For example, by the third program year, 90 percent of physicians in CPC+ practices reported patients had after-hours access to clinical staff with real-time access to electronic health records, compared with 80 percent of physicians in comparison practices (Swankoski et al. 2022). CPC+ practices were also required to increase the delivery of short-term, episodic care management, which involved timely outreach to patients after a hospital or emergency department discharge. By the third program year, a higher proportion of physicians in CPC+ than in comparison practices provided timely follow-up after emergency department visits and hospitalizations (Swankoski et al. 2022). Beneficiaries in Track 2 CPC+ practices were also more likely than those in comparison practices to report receiving timely follow-up after hospitalizations (CMS 2019b). As the model requires, CPC+ practices provided longitudinal care management services, though to only a relatively small percentage of high-risk patients (Swankoski et al. 2022). Stronger care management could lead to fewer exacerbations of the patient's underlying illness and, consequently, reduce the incidence of acute hospitalizations. Each process improvement could help strengthen primary care among CPC+ practices and reduce acute hospitalizations and less complex admissions, in particular.

Our findings also indicate meaningful differences in the impact of CPC+ on types of acute hospital admissions across subgroups. In particular, we saw non-SSP practices reduced less complex admissions, with some evidence that SSP practices influenced more complex admissions. These findings suggest there might be some positive interaction between incentives and supports CPC+ and SSP initiatives offer in being able to affect higher-complexity acute hospital care. We also saw that impacts were concentrated among independent practices, with no effects among the group of practices that were hospital- or system-owned. Differences in administration and financial incentives might explain why independent CPC+ practices were able to reduce acute hospitalizations, but hospital- or system-owned practices were not. Prior research indicates independent practices in CPC+ were more nimble than hospital- or system-owned practices and less likely to have the layers of internal bureaucracy that practices must navigate before implementing concrete steps to respond to incentives (Swankoski et al. 2022). It is also likely that hospital- or system-owned practices were not as aggressive about reducing hospital use because of the competing incentives they face; reducing hospital admissions would reduce revenues for medical and surgical specialists, and, for systems with hospitals, would directly reduce hospital revenues. Finally, our subgroup results showed beneficiaries in the highest quartile of the HCC score distribution had the largest reductions in hospital use. This finding aligns with additional room for improvements among these beneficiaries, who are more likely to have hospital use.

Our broad findings suggest CPC+ did not meaningfully affect more complex admissions, particularly surgical admissions, for Medicare beneficiaries, over the course of CPC+ (although there was evidence SSP practices had an effect in the last two program years). There are several possibilities that could explain these findings. First, CPC+ incentivized primary care practitioners but not surgical or medical

specialists. Although care coordination between primary and other providers is necessary for effective care management, surgical and medical specialists might require incentives to engage in behaviors that reduce surgical admissions. Second, primary care practices might require stronger, tailored learning supports and larger incentives to reduce higher acuity hospitalizations (Peikes et al. 2020). For example, some CPC+ practices indicated they could not provide longitudinal care management services to all patients who might benefit and noted that care managers had competing responsibilities that limited how much time they could spend providing longitudinal care management (Swankoski et al. 2022). Third, there might be a ceiling on the extent that high quality primary care can reduce more complex hospitalizations; it is possible that by delineating surgical versus medical hospitalizations (roughly), we can capture that ceiling.

5.J.5. Limitations

We note several limitations to this study. First, as with all our impact analyses, practices were not randomly assigned to CPC+ versus comparison group status. Therefore, differences in unobservable characteristics between CPC+ and comparison practices could influence outcomes, independently from the effects of the CPC+ model. Second, our measures for hospitalization severity are based on classification of claims data: electronic health records, which are more detailed, could provide more nuanced classifications of severity and potentially different results. Third, with the MS-DRG system, we cannot distinguish which hospitalizations with an MCC or a CC had a complication during the admission versus a comorbidity present at the time of admission (or both). This differentiation could be useful for understanding the mechanisms by which primary care shapes hospitalizations. Fourth, although we include regional COVID-19-related controls in our models, we cannot rule out that our estimates reflect differences between CPC+ and comparison regions in types of hospitalizations caused by the pandemic in PY 4 and PY 5. At the same time, the emergence of effects in PY 3 (and earlier for some outcomes), which was before the COVID-19 pandemic, somewhat alleviates this concern. Finally, the generalizability of our findings to other large-scale initiatives may be limited because CPC+ was tested in the regions, payers, and practices that volunteered to participate and were selected by CMS. Furthermore, given the flexibility of the CPC+ model, another set of practices might have transformed care differently, leading to different results.

5.J.6. Conclusion

Our findings indicate that large-scale initiatives to improve access to quality primary care, such as CPC+, hold promise for reducing lower-severity hospital admissions. However, such initiatives might have a limited impact on all-cause acute hospital expenditures. Together, our findings can help explain why CPC+ did not generate savings for Medicare: the types of hospital admissions CPC+ is reducing by are among the least costly admissions, and only a subset of practices (for example, independent practices) were able to achieve these reductions. More research about the contributions of specific care delivery changes CPC+ induces and how these processes differed across practice types could help explain our findings and shape future primary care models.

5.K. Effects of new enhanced payments

Key takeaways

We hypothesized that higher-than-average new enhanced payments per primary care practitioner (PCP) would lead to practices hiring more care managers and making greater improvements in care processes, specifically: providing more longitudinal care management (LCM), more episodic care management (ECM), more intense care management support, and more effective communication among providers.

We found that higher-than-average new enhanced payments per PCP were associated with:

- Hiring more care managers between Program Year (PY) 1 and PY 2, and
- Increasing the prevalence of episodic care management from PY 1 to PY 2.

But higher-than-average new enhanced payments per PCP were not significantly associated with increasing the prevalence of longitudinal care management, the intensity of care management, or the efficiency of providers' communication.

When we analyzed changes in episodic care management from PY 1 to PY 3, we found no association with new enhanced payments per PCP. This is possibly because practices with lower payments were "catching up" to practices with higher payments between PY 2 and PY 3. Because of limitations in our study, our results are best interpreted as suggestive evidence.

5.K.1. Introduction

New enhanced payments were a key CPC+ support because they allowed practices to invest in the infrastructure, staffing, and training necessary to deliver the five Comprehensive Primary Care Functions. The amount of new enhanced payments practices received from CMS (and, to a lesser degree, from payer partners) for CPC+ varied considerably, largely because of variations in the number and illness severity of patients attributed to the practice by Medicare fee-for-service (FFS) and other participating payers. Over the course of CPC+, practices made many changes to primary care processes, but it is unclear whether these changes depended on the size of enhanced payments practices received, or whether practices were able to make similar changes with below-average additional resources. In this analysis, we tested whether CPC+ practices that received higher-than-average new enhanced payments per PCP made more changes in care process measures in the first years of the intervention.

5.K.2. Hypotheses

During PY 1 (2017) deep-dive interviews, almost all practices reported that they spent their enhanced payments on care managers or coordinators.[120] Based on this, we hypothesized that:

H1. Higher-than-average new enhanced payments per PCP allowed practices to hire more care managers. This, in theory, might allow CPC+ practices to increase the following primary care processes:

[120] Some other practices reported using CPC+ payments to add other types of staff (e.g., behavioral health therapists, pharmacists, and social workers) and investments (e.g., expanded office hours); however, these were much less common across the practices interviewed.

H2. The prevalence of high-risk patients receiving LCM support

H3. The prevalence of patients receiving ECM

H4. The intensity of care management support

H5. Efficacy of providers' communication with one another about patients' care

We tested each hypothesis by analyzing whether practices that received more new enhanced payment amounts per PCP had larger changes in care manager staffing and primary care processes.[121]

5.K.3. Methods

To test the five hypotheses, we used regression models with practice self-reported changes in care process measures as the outcomes and a measure of new enhanced payment amounts per PCP as the main regressor, and we controlled for important practice characteristics. We used the sample of all 2017 Starter CPC+ practices that had outcome measures available in PY 1 (2017) and PY 2 (2018). The rest of this section provides more details on process-of-care outcome measures, the new enhanced payment per PCP measure, the model, outcome measures, and limitations to our analysis.

A. Process-of-care outcome measures

Because most care process changes happened between PY 1 and PY 2, our primary model analyzed changes in outcomes from PY 1 to PY 2. As an additional test, we also assessed changes in outcomes from PY 2 to PY 3.[122]

We identified outcome measures for the five hypotheses using CPC+ Practice Survey data and practice-reported care delivery data submitted to CMS (which we refer to as CDR [care delivery requirement] data). We identified items to include based on our review of all related questions that (1) were consistently asked over time, and (2) had a scale that indicated increasing intensity. For PY 1 CDR-based measures, we used the first quarter practices reported data on that measure (quarter 1 [Q1] or Q2),[123] and for the PY 2 and PY 3 CDR data, we used the last quarter practices reported data (Q3 or Q4).

For each variable we included in the care process measure, we rescaled it to be "pseudo-continuous" with values between 0 and 1. The one exception was the number of care managers per PCP measure, which we did not rescale to keep it in the units of full time equivalent (FTE) per PCP. For outcomes that included more than one variable, we calculated composite scores by taking the mean across all survey items included in the activity for each practice. Table 5.K.1 provides the source variables for each measure we used to test the hypotheses.

[121] For ease, throughout the rest of this Appendix, we refer to the collective outcomes for the five hypotheses as care process measures.

[122] Since most of the changes in process measures happened in the first three program years, we decided not to extend this analysis to look at changes through PY 4 or PY 5.

[123] The one exception to this is for the prevalence of the LCM measure, where we used the PY 1 Q2 values. This was because of some known data quality issues with the PY 1 Q1 values.

Table 5.K.1. Process of care outcome measures

Outcome measure	Source variables [data source, years available]
H1. Number of care managers per PCP	Number of care managers [PS, PY 1–PY 3] divided by number of PCPs [rosters, PY 1–PY 3]
H2. Prevalence of ECM	Average of 0-1 scaled variables: Portion of patients for which there is follow-up after ED [PS, PY 1–PY 3] Portion of patients for which there is follow-up after hospitalizations [PS, PY 1–PY 3]
H3. Prevalence of LCM	Number of patients under LCM [CDR – 3 years] divided by number of patients [CDR, PY 1–PY 3]
H4. Intensity of CM services	Average of 0-1 scaled variables: 1. Indicator for use of care plans [CDR, PY 1–PY 3] 2. Scale for documentation and integration of care plans [CDR, PY 1–PY 3] 3. Fraction of entities with access to care plans (out of five) [CDR, PY 1 – PY 2] 4. Indicator for whether any staff are responsible for developing care plans [CDR, PY 1–PY 3] 5. Development process of care plans [PS, PY 1–PY 2] 6. Frequency of sharing of care plans [PS, PY 1–PY 2] 7. Frequency of encouraging patients to choose goals [CDR, PY 1–PY 2] 8. Frequency of including caregivers in goal setting [CDR, PY 1–PY 2] 9. Frequency of measuring patients' skills and progress [CDR, PY 1–PY 2] 10. Frequency of providing self-management support [CDR, PY 1–PY 2] 11. Indicator for whether there is systematic identification of patients for SM support [CDR, PY 1–PY 2] 12. Number of conditions self-management is offered for [CDR, PY 1–PY 2] 13. Extent of self-management support (e.g., whether offered by trained professional) [PS, PY 1–PY 2]
	14. Frequency of updates to inventory of social service resources [CDR, PY 1–PY 2] 15. Extent of screening for unmet social needs [CDR, PY 1–PY 2] 16. Integration of social service resources and EHR [CDR, PY 1–PY 3] 17. Portion of patients assessment of social support needs is done for [PS, PY 1–PY 3] 18. Extent of linking patients to supportive community-based resources [PS, PY 1–PY 3] 19. Indicator for care manager on site [PS, PY 1–PY 3] 20. Indicator for care manager having a clinical background [PS, PY 1–PY 3] 21. Extent of pre-visit planning [PS, PY 1–PY 2]
H5. Communication with specialists	Average of 0-1 scaled variables: Number of specialties practice has care compact/collaborative agreement with [CDR, PY 1–PY 3] Frequency of timely receipt of info [PS, PY 1–PY 3] Number of specialist groups the practice has written agreements with [PS, PY 1–PY 3]

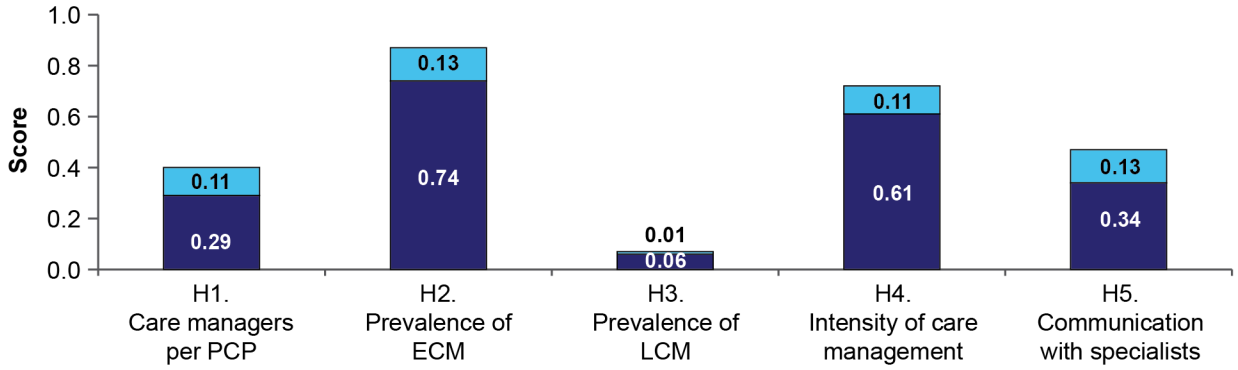
CDR = data practices report to CMS on their care delivery requirement activities; ED = emergency department; ECM = episodic care management; EHR = electronic health record; H = hypothesis; LCM = longitudinal care management; PY = Program Year; PS = Practice Survey.

Figure 5.K.1 shows the average level in PY 1 and the average change from PY 1 to PY 2 for each care process measure. All measures increased by at least 17 percent from PY 1 to PY 2.

- In PY 1, practices had on average 0.29 FTE care managers per PCP. This increased by 0.11 (or 37 percent) in PY 2.
- Practices had an average score of 0.74 for the fraction of discharges under ECM in PY 1, and this increased by 0.13 (or 18 percent) from PY 1 to PY 2.

- Practices had on average 6 percent of their patients under LCM (across all risk tiers) in PY 1, and this increased by 1 percentage point (or 17 percent) from PY 1 to PY 2.
- Practices had an average intensity of care management score of 0.61 in PY 1, and this increased by 0.11 (or 18 percent) on average in PY 2.
- Finally, practices had an average score of 0.34 for communication with specialists in PY 1 which increased by 0.13 (or 38 percent) in PY 2.

Figure 5.K.1. Changes in care process outcomes from PY 1 to PY 2 among CPC+ 2017 Starter Track 1 and Track 2 practices



Source: Mathematica's analysis of 2017 and 2018 CPC+ Practice Survey; and 2017 Q1–Q2, and 2018 Q3–Q4 CPC+ Practice Portal data.

Note: All care process outcomes, except for care managers per PCP, are on a scale of 0 to 1. Care managers per PCP is in FTEs per PCP units.

ECM = episodic care management; H = hypothesis; LCM = longitudinal care management; PCP = primary care practitioner; PY = Program Year.

B. New enhanced payment amount per PCP

We constructed the new enhanced payment amount per PCP as:

- PY 1 Medicare care management fees (CMFs), plus
- PY 1 CMFs from other payers who indicated they provided additional enhanced payments to support transformation compared to 2016, plus
- PY 1 imputed Performance-based Incentive Payments (PBIPs)[124]

Table 5.K.1 shows the average and standard deviation for each component of the new enhanced payment measure across CPC+ practices. Medicare CMFs are both the largest in size (with an average of $157,949 per practice) and have the largest variation (with a standard deviation of $164,022). CMFs from other papers are small on average ($11,857) but have a large degree of variation (with a standard deviation of $89,266). In addition, the variation in CMFs from other payers isn't driven primarily by practice size (which drives much of the variation in Medicare CMFs and imputed PBIPs), but by whether the practice contracts with payers that participate substantially in CPC+. The imputed PBIPs are the smallest on

[124] Imputed PBIPs were calculated as the PBIP the practice received at the beginning of 2017 multiplied by the average portion retained within their track. We used imputed PBIPs because the actual PBIPs reflect practice performance. While not directly linked to practice survey and CDR practice care process measures, they are determined by the Patient Experience of Care Survey, which are likely influenced by practices' care processes.

average ($7,840 per practice) and have the smallest amount of variation (with a standard deviation of $12,444).

Table 5.K.1. Summary statistics of the components of the new enhanced payments measure

	Average	Standard deviation
Medicare CMFs	157,949	164,022
CMFs from other payers	11,857	89,266
Imputed PBIPs	7,840	12,444
New enhanced payments	**177,645**	**203,840**

Source: Mathematica's analysis of PY 1 data on CPC+ payments provided by CMS, PY 1 practice-reported financial data submitted to CMS, and PY 1 CPC+ Payer Survey data.

Note: New enhanced payments were calculated as PY 1 Medicare CMFs plus PY 1 CMFs from other payers who indicated they provided additional enhanced payments to support transformation compared to 2016, plus imputed PBIPs (PBIP practice received at beginning of 2017 multiplied by the average portion retained within their track).

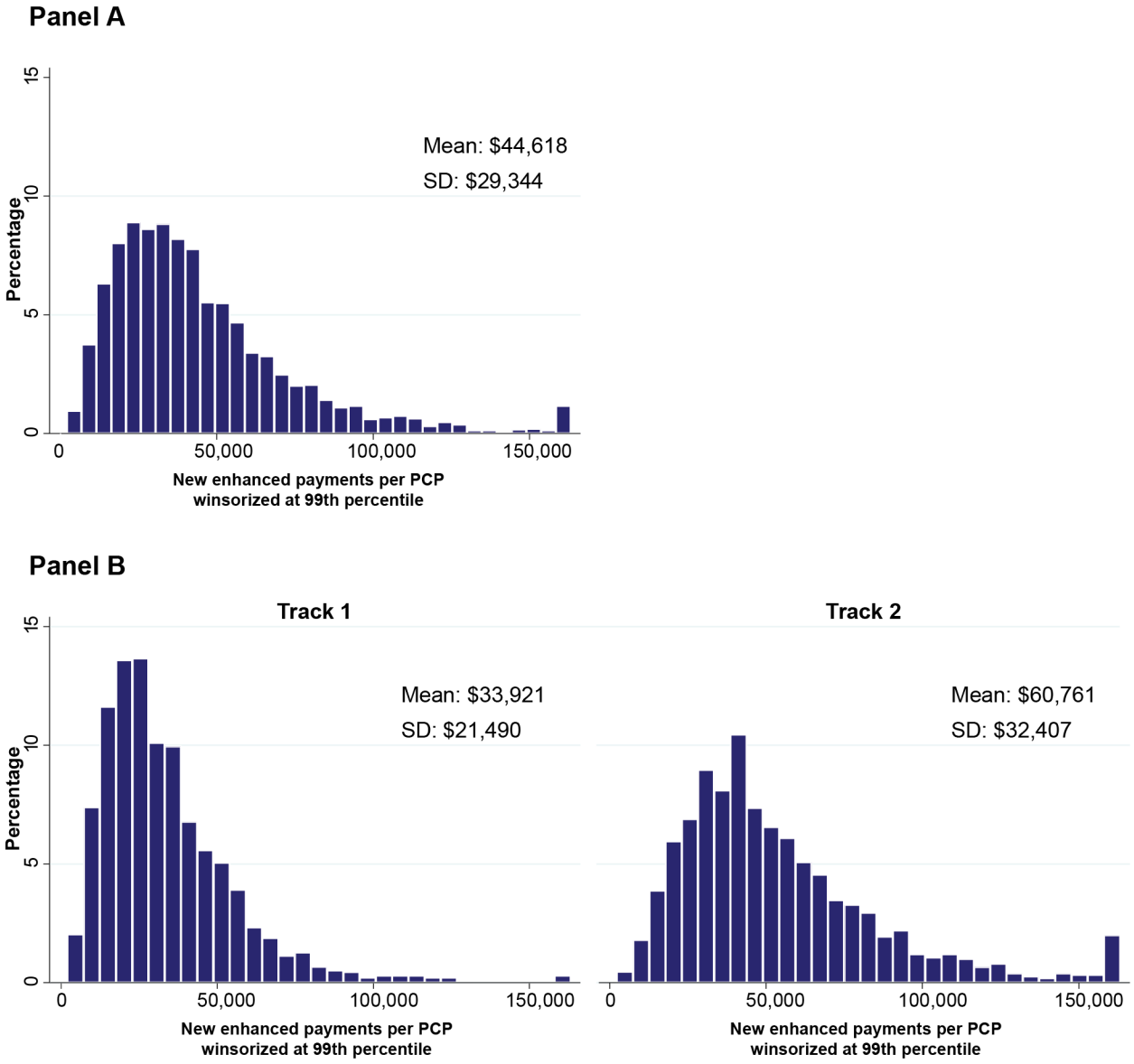
CMF = care management fee; PBIP = Performance-based Incentive Payment; PY = Program Year.

After creating new enhanced payments as the sum of the three components, we divided the sum by the number of PCPs in the practice, to create a per-PCP measure. There were some large outliers, so we winsorized this measure at the 99th percentile by setting values that were larger than the 99th percentile to the 99th percentile. Figure 5.K.2 shows the distribution of the winsorized new enhanced payments per PCP for the full sample and by track.

For the full sample, the average new enhanced payment amount per PCP was $44,618, with a standard deviation of $29,344. Split by track, we find that Track 1 practices had much lower average payments ($33,921 for Track 1 versus $60,761 for Track 2) and lower variation in payments (a standard deviation of $21,490 for Track 1 versus $32,407 for Track 2).

To estimate the effect of the average amount of new enhanced payments per PCP, we divided the new enhanced payment amount per PCP measure by its average ($44,618, which we refer to from now on as $45K).

Figure 5.K.2. Distribution of new enhanced payments per PCP in PY 1 among CPC+ 2017 Starter practices



Source: Mathematica's analysis of PY 1 data on CPC+ payments provided by CMS, PY 1 practice-reported financial data submitted to CMS, PY 1 CPC+ Payer Survey data, and March 2017 practice-reported roster data submitted to CMS.

Note: New enhanced payments were calculated as PY 1 Medicare CMFs plus PY 1 CMFs from other payers who indicated they provided additional enhanced payments to support transformation compared to 2016, plus imputed PBIPs (PBIP practice received at beginning of 2017 multiplied by the average portion retained within their track) per primary care practitioner in the practice (based on the March 2017 roster). The measure was winsorized at the 99th percentile, which means all values above the 99th percentile were replaced by the 99th percentile value.

CMF = care management fee; PBIP = Performance-based Incentive Payment; PCP = primary care practitioner; PY = Program Year; SD = standard deviation.

C. Model

Our main model regressed changes in care process measures from PY 1 to PY 2 on the new enhanced payment amount per PCP practices received in PY 1. We also included the level of the care process measure in PY 1, as well as beneficiary, practice, and geographic controls.

- Beneficiary demographic characteristic controls included average beneficiary age and the proportions for each of the following categories: age, race, sex, original entitlement reason, and dual eligibility. Risk controls included average hierarchical condition category (HCC) score and proportions for the following categories: Tier 4, and Tier 5 beneficiaries (defined based on payment attribution methodology).
- Geographic characteristic controls included an indicator for a primary care health professionals shortage area, indicators for the practice being rural or suburban, the percentage of county-level poverty in 2014, and the percentage of county-level Medicare Advantage enrollees/eligible beneficiaries in 2015.
- Practice characteristic controls included an indicator for being in Track 2, an indicator for ownership by a hospital, an indicator for ownership by a health system or hospital, categorical counts of PCPs, a quadratic in the number of PCPs, categorical counts of providers, a quadratic in the number of providers, a quadratic in the number of patients, an indicator for participation in the Shared Savings Program as of January 1, 2017, an indicator for participation in the downside risk tracks (2 or 3) in the Shared Savings Program as of January 1, 2017, an indicator for experience with a prior transformation initiative (e.g., participated in the Multi-Payer Advanced Primary Care Practice model, medical home recognition, or participated in CPC Classic), an indicator for whether the practice has providers from multiple specialties, an indicator for whether practice has any nurse practitioners or physician assistants, an indicator for being majority native, and the percentage of charges that are primary care.

Although we controlled for the size of the practice in terms of the number of PCPs and the total number of patients, we did not control for the number of attributed beneficiaries because the number of beneficiaries per PCP accounted for a large portion of the variation in the new enhanced payment amount per PCP measure. Track also contributed to a large portion of the variation in the new enhanced payment amount per PCP measure, but because there are other important differences in the intervention by track, like care delivery requirements, we included an indicator for Track 2 as a control in our main model.

D. Limitations

One limitation of this analysis is that we did not have measures of care processes prior to the start of CPC+ in 2017. We used the earliest available data from the PY 1 Practice Survey (administered from May 2017 through September 2017) and PY 1 Q1 and Q2 responses from the CDR. However, practices may have already made changes to their care processes prior to our measurements, which could understate the relationship between new enhanced payments and the changes in care processes in our models. Also, changes in these self-reported measures may not accurately reflect the changes practices made in care processes, for example, due to some of the measures not having much room for improvement at the beginning of the intervention and some practices not interpreting the questions correctly.

In addition, new enhanced payment amounts per PCP are not randomly assigned, and practices that received higher-than-average new enhanced payment amounts per PCP may have been different in other ways that could affect the care process changes they made. A large portion (38 percent) of the variation in the new enhanced payment amount per PCP measure comes from the number of beneficiaries per PCP in the practice, which is likely attributable to differences in the percentage of practices' patient populations that are Medicare beneficiaries. If practices that have a larger portion of Medicare beneficiaries in their patient panels differ in other ways that contribute to changes in care processes that practices make, our estimates could be biased. For example, such practices may have more interest in primary care transformations that are aimed specifically at their patient population than practices with a smaller portion of Medicare beneficiaries among their patients. Although we controlled for many practice characteristics, we did not control for the main source of variation that comes from the portion of Medicare beneficiaries in the patient panel.

5.K.4. Results

Higher new enhanced payment amounts per PCP were positively and significantly associated only with changes in the number of care managers per PCP and the prevalence of ECM (Table 5.K.2).

Table 5.K.2 reports the results from regressions on each different care process outcome (row). The first column provides the average change in the care process outcome from PY 1 to PY 2, then gives the coefficient from the regression of the change in the care process outcome on the new enhanced payment amount per PCP (that is, the effect of the average amount of new enhanced payments per PCP on the outcome), and finally calculates the percentage of the change in the outcome that can be explained by the effect of the new enhanced payment amount per PCP (that is, the average change in care process outcome divided by the estimated coefficient).

H1. Care Managers per PCP. We estimated that practices that received $45K more than the average enhanced payment per PCP increased the number of care managers per PCP by 0.075 FTEs more than practices with the average enhanced payment per PCP ($p < 0.01$). This suggests that 71 percent of the average change in care managers per PCP from PY 1 to PY 2 (0.11) can be explained by the new enhanced payments that practices received in PY 1.

H2. Prevalence of ECM. Practices that received $45K more than the average new enhanced payments per PCP significantly increased the prevalence of ECM score by 0.018 ($p < 0.01$), or 14 percent of the average change in the prevalence of ECM from PY 1 to PY 2.

H3. Prevalence of LCM. We found no statistically significant effect of higher-than-average new enhanced payments per PCP on the change in the prevalence of LCM.

H4. Intensity of care management. We found no statistically significant effect of higher-than-average new enhanced payments per PCP on the change in the intensity of care management.

H5. Communication with specialists. Practices that received $45K more than the average new enhanced payment per PCP decreased their communication with specialists score by 0.02 ($p < 0.10$). This suggests that practices that received higher-than-average new enhanced payments per PCP made fewer improvements in the communication with specialists. The mechanism for a negative causal relationship is unclear, which suggests that this result may be driven by bias or may be spurious and due to multiple hypothesis testing.

Table 5.K.2. Effects of new enhanced payments on change in care process outcomes from PY 1 to PY 2

	Average change in outcome PY 1 to PY 2	Coefficient (SE)	% of change in outcome explained
H1. Care managers per PCP	0.11	0.075*** (0.019)	71%
H2. Prevalence of ECM	0.13	0.018*** (0.007)	14%
H3. Prevalence of LCM	0.01	0.002 (0.005)	31%
H4. Intensity of care management	0.11	0.004 (0.003)	3%
H5. Communication with specialists	0.13	-0.020*** (0.006)	NA

Source: Mathematica's analysis of 2017 and 2018 CPC+ Practice Survey; 2017 Q1–Q2 and 2018 Q3–Q4 CPC+ Practice Portal data, PY 1 data on CPC+ payments provided by CMS, PY 1 practice-reported financial data submitted to CMS, PY 1 CPC+ Payer Survey data, and March 2017 practice-reported roster data submitted to CMS.

Note: Outcome measures ranged from 0 to 1, except for care managers per PCP, which was in FTE per PCP units. Each row represents a separate regression model on a different outcome. Each model controls for practice, geographic, and beneficiary characteristics at the practice level. See methods section for list of controls.

[a] We calculated the percentage of the effect explained by taking [estimated impact of average new enhanced payment/average change in outcome]*100. In some cases, the estimated impact was negative and could not explain the positive change in the outcome. In those cases, the percentage of the change in the outcome explained is shown as NA.

*/**/*** Significantly different from zero at the 0.10/0.05/0.01 level, two-tailed test.

ECM = episodic care management; FTE = full-time equivalent; LCM = longitudinal care management; NA = not applicable; PCP = primary care practitioner; PY = Program Year; SE = standard error.

Results were generally not sensitive to changes in the controls; however, when we didn't include an indicator for Track 2, we found positive significant effects of higher-than-average new enhanced payments per PCP on intensity of care and communication with specialists (not shown). This suggests that those increases resulted from practices being in Track 2 (perhaps due to the more stringent care delivery requirements) and not because they had received higher-than-average new enhanced payments per PCP.

There was no evidence that effects differed significantly for subgroups. Results from restricting to Track 1 practices were similar to those from restricting to Track 2 practices, and results from restricting to hospital- or system-owned practices were similar to those from restricting to independent practices (not shown).

We found no additional effect of new enhanced payment amounts per PCP on changes to the number of care managers per PCP between PY 2 and PY 3; and we found a negative effect on changes to the prevalence of ECM between PY 2 and PY 3, suggesting that payments were important for early changes but not for later ones.

Although most of the total changes in the care process measures during the intervention occurred between PY 1 and PY 2, some additional improvements occurred between PY 2 and PY 3. To see how more new enhanced payments per PCP affected outcomes over a longer period, we looked at changes in the outcomes from PY 2 to PY 3 as well.

This change in time period resulted in a couple of additional changes in our analysis:

- For the intensity of care management measure, there were fewer items from the practice survey and CDR that were consistently available from PY 2 to PY 3 (8 versus 21 for the PY 1 to PY 2 change).
- For all measures, there were approximately 100 fewer practices that had responses (due to practices leaving CPC+ and not responding to the Practice Survey and CDR portal questions).

Because of these changes, we also reassessed the impact of new enhanced payments on the PY 1 to PY 2 change in care process measures using the updated sample and measures.

Table 5.K.3 reports the results on the effect of having higher-than-average new enhanced payments per PCP on changes in care process outcomes from PY 1 to PY 2 and PY 2 to PY 3. The results for the PY 1 to PY 2 changes are very similar to those reported in Table 5.K.2. When we looked at the change from PY 2 to PY 3, we found no effect of higher-than-average new enhanced payments per PCP on care managers per PCP. This is perhaps not surprising given there was no additional increase in care managers per PCP from PY 2 to PY 3. We also found a significant *negative* effect of higher-than-average new enhanced payments per PCP on the prevalence of ECM from PY 2 to PY 3, which was of the same magnitude (but of a different sign) than the PY 1 to PY 2 effect. This result suggests that practices that received higher-than-average new enhanced payments per PCP made early improvements in the prevalence of ECM, but then practices with lower-than-average enhanced payments "caught up" by PY 3.

Table 5.K.3. Effects of new enhanced payments on change in care process outcomes from PY 1 to PY 3

	Change in outcome PY 1-PY 2	Impact on PY 1-PY 2 change – Coefficient	Change in outcome PY 2-PY 3	Impact on PY 2-PY 3 change – Coefficient
H1. Care managers per PCP	0.11	0.075***	0.00	0.003
H2. Prevalence of ECM	0.13	0.023***	0.06	-0.023***
H3. Prevalence of LCM	0.01	0.006	-0.01a	-0.003
H4. Intensity of care management	0.10	-0.004*	0.03	-0.003
H5. Communication with specialists	0.13	-0.011*	0.08	0.008

Source: Mathematica's analysis of 2017, 2018, and 2019 CPC+ Practice Surveys; and 2017 Q1–Q2, 2018 Q3–Q4, and 2019 Q3 CPC+ Practice Portal data. Program Year (PY) 1 is 2017, PY 2 is 2018, and PY 3 is 2019. PY 1 practice-reported financial data submitted to CMS, PY 1 CPC+ Payer Survey data, and March 2017 practice-reported roster data submitted to CMS.

Note: Outcome measures ranged from 0 to 1, except for care managers per PCP, which was in FTE per PCP. Each row represents a separate regression model on a different outcome. Each model controls for practice, geographic, and beneficiary characteristics at the practice level. See methods section for list of controls.

[a] The mean value of the fraction of patients under LCM decreased from PY 2 to PY 3, but the median increased across all three years from 0.004 in PY 1 to 0.023 in PY 2 to 0.030 in PY 3.

*/**/*** Significantly different from zero at the 0.10/0.05/0.01 level, two-tailed test.

ECM = episodic care management; FTE = full-time equivalent; LCM = longitudinal care management; PCP = primary care practitioner; PY = Program Year.

5.K.5. Discussion

We found that practices that received higher-than-average new enhanced payments per PCP at the start of CPC+ hired more care managers per PCP and made more improvements to the prevalence of ECM between PY 1 and PY 2. However, when we assessed the impact of higher-than-average new enhanced payments on changes between PY 2 and PY 3, we found no additional effect on care managers per PCP. Additionally, the impact on the prevalence of ECM was negative and similar in magnitude to the increase from PY 1 to PY 2, suggesting that practices with lower payments were able to "catch up" to the progress practices with higher payments made by PY 3.

This suggests that practices with more new enhanced payments per PCP were able to hire more care managers per PCP. However, this did not translate into additional measured care process improvements in the longer term. Possible reasons for this include:

1. First, practices may have made improvements in the care process measures prior to our first measurement. If practices with more new enhanced payments per PCP were more likely to make these early changes (and less likely to make additional improvements), we might not identify an effect.
2. Second, some of the measures—particularly the measure of the prevalence of ECM—had little room for improvement after PY 2. We could not measure improvement in the prevalence of ECM beyond practices reporting they followed up with "most or all" of their patients within one week of an ED visit and three days of hospital discharge. It is possible practices that received higher-than-average new enhanced payments per PCP were able to make more subsequent improvements to the prevalence of ECM that we cannot measure.
3. Third, practices may not have responded to the Practice Survey and CDR questions correctly and consistently over time.
4. Fourth, our results could be biased due to associations between the new enhanced payments per PCP and other practice characteristics that determine changes in care processes.

5.K.6. Conclusion

New enhanced payments are a key CPC+ support that the model provided to practices to make primary care transformation. Although we found that higher-than-average new enhanced payments per PCP led practices to hire more care managers, we did not see strong sustained effects on care process changes related to care management and communication with specialists. This could be because of bias in our estimates due to measurement issues with these processes over time or because the amount of payments practices received were not randomized and instead were directly related to the number of Medicare beneficiaries they had. Or it could suggest that even practices that received below-average enhanced payments per PCP were able to eventually achieve improvements in care process measures.

Given the limitations of our study, further investigation is needed to better understand whether the size of new enhanced payments affects the amount of care process changes practices make. To do so, researchers and model evaluators need better measures of important care process activities and accurate measurements of them over time, including measurements prior to the start of the model. Survey items should be designed to measure care process activities explicitly and concretely, with ample room for growth. In addition to better measures of activities, if policymakers want to be able to understand whether more new enhanced payments lead to more care process changes (and better ultimate beneficiary outcomes), future innovation models should explicitly test this, for example, through random assignment of payments for participating practices.

5.L. Patient Choices Pathway analysis

Key takeaway

We tested whether improved access to primary care influenced patients' choices about where to seek care after symptoms developed for conditions that may be treated in a primary care setting. While access to primary care modestly improved among CPC+ practices during the intervention period, we find that the improvements that we can measure with survey data cannot explain the small effect of CPC+ on primary care substitutable emergency department (ED) visits.

5.L.1. Introduction

Over the first three program years, CPC+ reduced outpatient ED visits for Track 1 and 2 practices, relative to comparison practices. We found cumulative reductions of 7.2 visits per 1,000 beneficiaries (-1.5 percent; $p < 0.01$) among Track 1 practices and 7.2 visits per 1,000 beneficiaries (-1.5 percent; $p < 0.01$) among Track 2 practices (Peikes et al. 2021b). These results were consistent with previous research that found Comprehensive Primary Care (CPC) practices had 2 percent reductions in ED visits compared to comparison practices (Timmins et al. 2020).

Based on the CMS implementation guide and related literature, we identified two potential mechanisms through which reductions in ED visits might occur: (1) improved patient health status by better management of chronic illness and behavioral health conditions, and (2) improved access influencing patients' choices for where to seek care after symptoms develop for conditions that can be treated in a primary care setting (Peikes et al. 2021a).

The goal of this patient choices causal pathway analysis was to determine the extent to which CPC+ reduced ED outcomes through the second mechanism. Previous results from the Third Annual Report showed that, relative to comparison practices, CPC+ decreased primary care substitutable (PCS) outpatient ED visits (that is, ED visits for conditions that may be treated in a primary care setting) over the first three program years by 3.5 visits per 1,000 beneficiaries per year for Track 1 practices (-1.9%, $p = 0.002$) and 4.7 visits per 1,000 beneficiaries per year for Track 2 practices (-2.5%, $p < 0.001$) (Peikes et al. 2021b). These reductions in the PCS ED visit rates accounted for 48 to 64 percent of the reduction in overall outpatient ED visits, respectively. Based on these findings, it is possible that improvements in access may have driven about one-half of the reductions in outpatient ED visits.

In this memo, we investigate whether: (1) access to primary care improved over time among CPC+ practices, (2) CPC+ reduced PCS ED visits on weekdays and non-weekdays, and (3) greater changes in access were associated with fewer PCS ED visits among CPC+ practices over the first three program years.[125]

[125] Due to data constraints at the time this analysis was conducted, we only analyzed data through PY 3. As a follow-up, Appendix 5.O. Synthesis describes results from regressions of changes in promising care processes from PY 1 to PY 2 on changes in outcomes from PY 2 through PY 5.

5.L.2. Data, sample, and methods

We created primary measures of access using available CPC+ Practice Survey and Care Delivery Reporting (CDR) Portal data to calculate composite scores that ranged from 0 to 1 (see Table 5.L.1). These scores included an overall access composite measure and composite measures of three components of access: extended hours of operation, 24/7 access to the care team, and same- or next-day appointments. Additionally, we calculated several supplemental measures of access: percentage of primary care (PC) visits occurring on weekends, as well as the numbers of primary care practitioner (PCP) physician, PCP mid-level, and care manager full-time equivalents (FTEs) per 1,000 patients. Then, among the 2017 Starter CPC+ practices, we evaluated trends in composite score means from the Program Year (PY) 1 (2017, the earliest available CPC+ Practice Survey and CDR Portal data) to PY 2 (2018).[126]

Table 5.L.1. Source variables for access composite scores

Composite score type	Question	Data source
24/7 access to care	Patient after-hours access (24 hours, 7 days a week) to a physician, PA/NP, or nurse… Is not available or is limited to an answering machine; is available from a coverage arrangement (e.g., answering service) that does not offer a standardized communication protocol back to the practice for urgent problems; Is provided by a coverage arrangement (e.g., answering service) that shares necessary patient data with and provides a summary to the practice; OR Is available via the patient's choice of email or phone directly with the practice team or a practitioner who has real-time access to the patient's electronic medical record. (rescaled 0 through 1)	CPC+ Practice Survey
	Does a clinician or care team member from your practice site usually provide 24/7 coverage? Yes No	CDR
	Is 24/7 coverage provided with real-time access to your practice's EHR? Yes No	CDR
Expanded office hours	When patients need it, my practice is able to provide office visits during expanded hours on the weekend, evening, or early morning… Never Rarely Sometimes Often Always (rescaled 0 through 1)	CDR
Same-day or next-day appointments	Same-day appointments for patients who need them are available at this practice site for … None of this practice's patients. Some of this practice's patients. Many of this practice's patients. Most or all of this practice's patients. (rescaled 0 through 1)	CPC+ Practice Survey

[126] We also evaluated trends from PY 1 through PY 3 for the primary measures of access. Changes were driven by increases from PY 1 to PY 2. The overall access score increased from PY 2 to PY 3 by 2 percent (from 0.87 to 0.89) among CPC+ Track 1 practices and did not change for Track 2 CPC+ practices from PY 2–PY 3 (0.90 in PY 2 and PY 3). Examining changes from PY 1– PY 2 also informed our changes-in-changes regression analyses.

Composite score type	Question	Data source
	When patients need it, my practice is able to provide same- or next-day appointments… Never Rarely Sometimes Often Always (rescaled 0 through 1)	CDR
Overall CPC+ practice-reported access	All of the above	CPC+ Practice Survey, CDR
Percentage of primary care visits occurring on weekends	Percentage of all ambulatory primary care visits that occurred on weekends is calculated by dividing the count of primary care visits that occurred on weekends by the count of total ambulatory primary care visits (that occurred on weekdays or weekends) at the practice level and multiplying the result by 100.	Claims
PCP Physician Staffing	PCP physician FTEs per 1,000 patients is calculated by dividing the FTEs (from the Practice Survey) by the number of active patients in the practice (from the CDR) and multiplying the result by 1,000. PCP physicians include physicians and physician residents or fellows.	CPC+ Practice Survey, CDR
PCP Mid-levels Staffing	PCP mid-level FTEs per 1,000 patients is calculated by dividing the FTEs (from the Practice Survey) by the number of active patients in the practice (from the CDR) and multiplying the result by 1,000. PCP mid-levels include nurse practitioners (NPs), physician assistants (PAs), and clinical nurse specialists (CNSs).	CPC+ Practice Survey, CDR
Care Manager Staffing	Care Manager FTEs per 1,000 patients is calculated by dividing the FTEs (from the Practice Survey) by the number of active patients in the practice (from the CDR) and multiplying the result by 1,000.	CPC+ Practice Survey, CDR

Note: For all three staffing measures, we excluded outlier practices that reported only one or two patients in PY 1 (which are the practices that initially applied for CPC+ as a single practice and later split) or over 50,000 patients (three practices) in PY 1 and ensured practices had non-missing values in both years. The CPC+ Program Year (PY) 1 is 2017.

EHR = electronic health record; FTE = full-time equivalent; NP = nurse practitioner; PA = physician assistants; PCP = primary care practitioner; PY = program year.

We applied the New York University Emergency Department Algorithm (NYU EDA) to outpatient ED visits using primary diagnosis codes to identify PCS ED visits, which is the combination of non-emergent and emergent/primary care (PC) treatable ED visits as defined by the NYU EDA (Billings et al. 2000 and Johnston et al. 2017). We used Medicare claims from January 2016 through December 2019 to examine whether CPC+ led to fewer PCS ED visits on weekdays and non-weekdays over the first three program years using a practice fixed effects difference-in-differences (DD) model, which included 2017 Starter CPC+ practices as well as comparison practices. The model used weights accounting for beneficiary eligibility and clustered standard errors at the practice level. This model also included beneficiary covariates (for example, baseline [2016] characteristics including hierarchical condition category [HCC] score, new enrollee and dual eligibility status, and age, gender, race, original entitlement reason, and chronic condition categories) and an interaction of the cumulative intervention post-period (2017–2019) indicator with the treatment indicator.

We investigated whether changes in access from PY 1 to PY 2 were associated with changes in total PCS ED visit utilization from baseline to PY 3 (2019).[127] We estimated separate regressions for each access measure.[128] This "changes-in-changes" model was a practice-level model with practice-level clustered standard errors, which included only 2017 Starter CPC+ practices, separately by track, and weighted by practice size (baseline number of attributed beneficiaries). This model also controlled for the PY 1 access score and the baseline outcome rate, as well as other baseline beneficiary, practice, and geographic factors.[129]

Lastly, we calculated the percentage of the effect explained by the access component by taking $\left(\frac{(\text{estimated changes-in-changes impact}) * (\text{mean change in score from } 2017-2018)}{\text{PCS ED difference-in-differences effect coefficient}}\right) * 100$

Table 5.L.2 summarizes the data sources, years, and sample included in each of these analyses.

Table 5.L.2. Sample, data sources and years by analysis

Analysis	Sample	Data sources and years
Trends in access to primary care	2017 Starter CPC+ practices only	CPC+ Practice Survey: 2017 and 2018 CDR Portal: 2017 (Q1 or Q2) and 2018 (Q3 or Q4)
Difference-in-differences	2017 Starter CPC+ practices and comparison practices	Outcome: Medicare claims from January 2016 through December 2019a
Changes-in-changes	2017 Starter CPC+ practices only	Outcome: Changes in utilization from 2016 to 2019, using Medicare claims from January 2016 through December 2019a Key explanatory variable: Changes in access to primary care from 2017 to 2018.b Access is measured using CPC+ Practice Survey: 2017 and 2018 CDR Portal: 2017 (Q1 or Q2) and 2018 (Q3 or Q4)

[a] The CPC+ baseline period corresponds to 2016, Program Year (PY) 1 is 2017, PY 2 is 2018, and PY 3 is 2019.

[b] We examined changes in access to primary care from PY 1 to PY 2 to allow time for the practices' care delivery changes to impact the beneficiary service utilization outcomes through PY 3 (i.e., PCS ED outcomes in PY 3 informing practices' determination of access in PY 3).

[127] We examined changes in access to primary care from PY 1 to PY 2 to allow time for the practices' care delivery changes to impact the beneficiary service utilization outcomes through the PY 3, and to remove the possibility of reverse causality (i.e., PCS ED visit outcomes in PY 3 informing practices' determination of access in PY 3).

[128] We ran sensitivity models that controlled for the different components of access in the same model and generally found similar results.

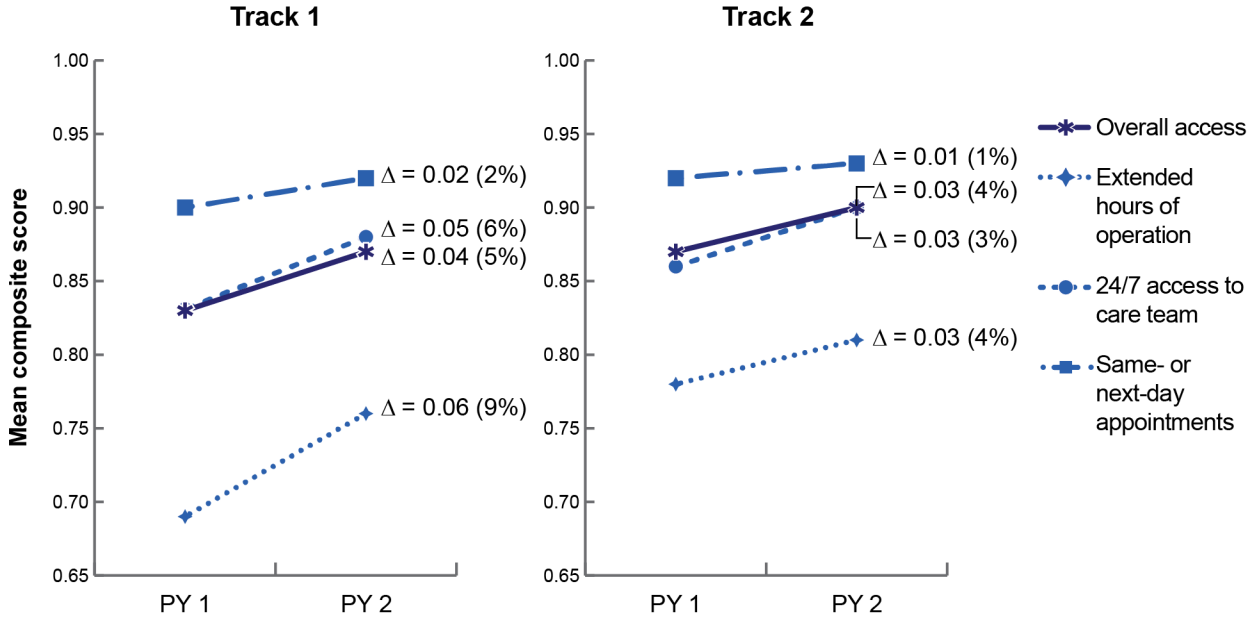
[129] We ran sensitivity models that varied the set of control variables we used and generally found similar results; starting with unadjusted results only controlling for the PY 1 access score and baseline outcome rate, we progressively added more control variables by category (i.e., baseline beneficiary demographics, baseline HCC scores and chronic conditions, baseline geographic characteristics, baseline practice characteristics [main model results], and change in beneficiary characteristics [not included in main model results]).

5.L.3. Results

A. Changes in access among CPC+ practices

In the ways we could measure access with our data sources, we observed improvements in primary measures of access from PY 1 to PY 2 among CPC+ practices. The mean overall access composite measure score increased by 5 percent for Track 1 practices and 3 percent for Track 2 practices, improving from the PY 1 scores of 0.83 and 0.87, respectively (Figure 5.L.1). Among the different components of the overall access score, the extended hours of operation component increased the most from PY 1 to PY 2, improving the Track 1 PY 1 mean score of 0.69 by 9 percent and the Track 2 PY 1 mean score of 0.78 by 4 percent (Figure 5.L.1). We observed mean 24/7 access to the care team component scores of 0.83 and 0.86 during PY 1 for Track 1 and Track 2 practices, respectively. The average scores for this component increased by 6 percent for Track 1 and by 4 percent for Track 2 practices between PY 1 and PY 2 (Figure 5.L.1). The same- or next-day appointments component of access was most topped out during PY 1, with mean scores of 0.90 and 0.92 for Track 1 and Track 2 practices, respectively; however, the average score for this component did improve by 2 percent (Track 1) and 1 percent (Track 2) between PY 1 and PY 2 (Figure 5.L.1). Overall, Track 2 practices had had less room for improvement with higher access scores during PY 1 and experienced lower change from PY 1 to PY 2 compared to Track 1 (Figure 5.L.1).

Figure 5.L.1. Trends over time in main components of access



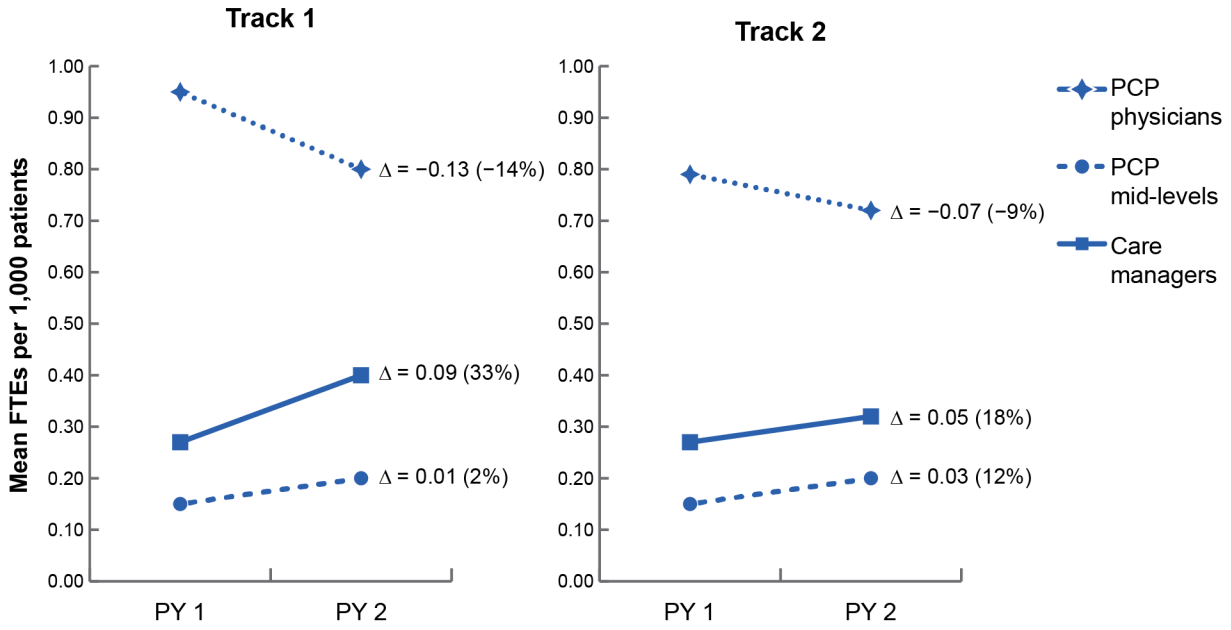
Source: Mathematica's analysis of 2017 and 2018 CPC+ Practice Survey, as well as 2017 Q1-Q2 and 2018 Q3-Q4 CPC+ Practice Portal data.

Note: Access scores could range from 0 to 1. All included variables were rescaled to 0 to 1 and then averaged to create the composite. PY 1 is 2017 and PY 2 is 2018.

PY = Program Year

Among the supplemental measures of access that we used as proxies for the points of contact with the practice, PCP mid-level and care manager staffing FTEs per 1,000 patients improved from PY 1–PY 2. PCP mid-level staffing FTEs per 1,000 patients increased by 2 percent for Track 1 and 12 percent for Track 2, and care manager staffing FTEs per 1,000 patients increased by 33 percent for Track 1 and 18 percent for Track 2 (Figure 5.L.2). For the other supplemental measures of access, PCP physician staffing FTEs per 1,000 patients and the percentage of PC visits occurring on weekends, we did not find large improvements; in fact, PCP physician staffing FTEs per 1,000 patients decreased from PY 1 to PY 2 (see Figures 5.L.2 and 5.L.3).

Figure 5.L.2. Trends over time in PCP physician, PCP mid-level, and care manager staffing supplemental measures of access

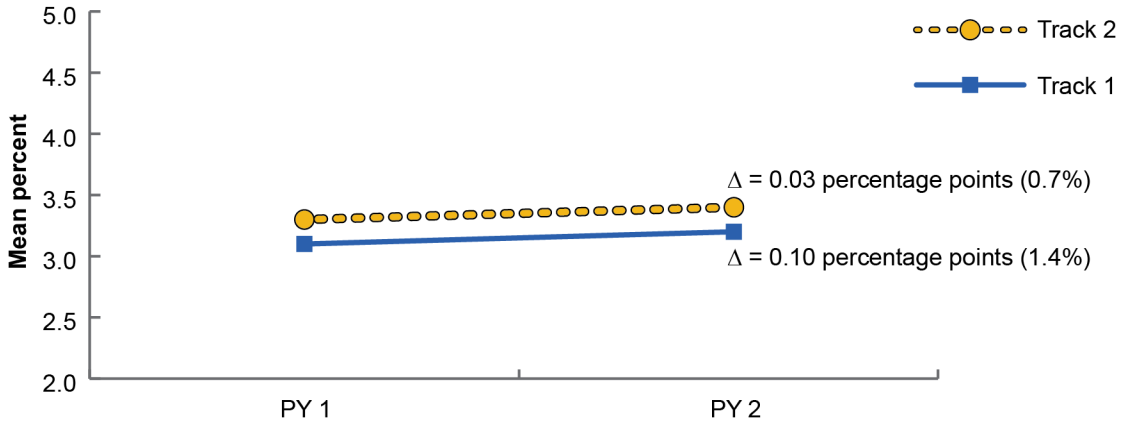


Source: Mathematica's analysis of 2017 and 2018 CPC+ Practice Survey, as well as 2017 Q1-Q2 and 2018 Q3-Q4 CPC+ Practice Portal data.

Note: We excluded outlier practices that reported only one or two patients in PY 1 (which are the practices that initially applied for CPC+ as a single practice and later split) or over 50,000 patients (three practices) in PY 1 and ensured practices had non-missing values in both years. PY 1 is 2017 and PY 2 is 2018.

FTE = full-time equivalent; PCP = primary care practitioner; PY = Program Year.

Figure 5.L.3. Trends over time in percentage of primary care visits occurring on weekends supplemental measure of access



Source: Mathematica's analysis of Medicare claims data from PY 1 (2017) to PY 2 (2018).
PY = Program Year

B. The effect of CPC+ on primary care substitutable ED visits

CPC+ decreased the weekday PCS ED visit rate relative to comparison practices for both tracks by 3.1 visits (-2.4 percent, $p < 0.001$). CPC+ also decreased the non-weekday PCS ED visit rate relative to comparison practices by 1.5 visits (-2.6 percent, $p < 0.001$) for Track 2 only (Table 5.L.3). Because more ED visits occur during the weekdays, impacts on weekday visits are larger in absolute terms, but they are also larger in relative terms (as a percentage of the comparison group mean) for Track 1. This may suggest that practices' access-related changes are happening on weekdays (for example, more same- or next-day appointments, rather than improved hours of operation on weekends). This is consistent with the CPC+ Exemplar study, which found that many exemplar practices increased same-day visits (Laird et al. 2022, Appendix 4C). These Track 1 findings were also consistent with previous research that found CPC practices had reductions in PCS ED visits compared to comparison practices, which was driven by reductions in weekday PCS ED visits (Timmins et al. 2020).

Table 5.L.3. Regression-adjusted means and estimated impact of CPC+ on outpatient ED visits and overall, weekday, and non-weekday primary care substitutable ED visits, for attributed Medicare FFS beneficiaries over the first three program years, Track 1 and Track 2 practices

	Track 1						Track 2					
Service use (per 1,000 beneficiaries per year)	CPC+ mean[a]	C mean[b]	Impact estimate[c] (SE)	Percentage impact	90% confidence interval	*p*-Value	CPC+ mean[a]	C mean[b]	Impact estimate[c] (SE)	Percentage impact	90% confidence interval	*p*-Value
Outpatient ED visits, including observation stays												
Baseline	493	498	NA	NA	NA	NA	492	492	NA	NA	NA	NA
PY 1 through 3	485	498	-7.2*** (2.3)	-1.5%	(-11.0, -3.4)	0.002	484	491	-7.3*** (2.4)	-1.5%	(-11.2, -3.3)	0.003
Primary care substitutable (PCS) outpatient ED visits, including observation stays [d]												
Baseline	192	195	NA	NA	NA	NA	191	192	NA	NA	NA	NA
PY 1 through 3	184	192	-3.5*** (1.1)	-1.9%	(-5.4, -1.7)	0.002	183	189	-4.7*** (1.1)	-2.5%	(-6.5, -2.8)	0.000
Weekday PCS ED visits, including observation stays [d]												
Baseline	132	134	NA	NA	NA	NA	131	132	NA	NA	NA	NA
PY 1 through 3	127	133	-3.1*** (0.8)	-2.4%	(-4.5, -1.7)	0.000	127	131	-3.1*** (0.9)	-2.4%	(-4.5, -1.7)	0.000
Non-weekday (weekends and holidays) PCS ED visits, including observation stays [d]												
Baseline	60	61	NA	NA	NA	NA	60	60	NA	NA	NA	NA
PY 1 through 3	57	59	-0.5 (0.5)	-0.8%	(-1.3, 0.3)	0.323	57	58	-1.5*** (0.5)	-2.6%	(-2.3, -0.8)	0.000

Source: Mathematica's analysis of Medicare claims data from January 2016 through December 2019. The CPC+ baseline period corresponds to 2016, Program Year (PY) 1 is 2017, PY 2 is 2018, and PY 3 is 2019.

Note: This table indicates which estimates are statistically significant; when we interpret evidence, we combine evidence from the magnitude of the effect, the p-values, findings on related outcomes, subgroups, sensitivity tests, and other data sources about model implementation. These results are from the CPC+ third annual report, and they provide context for our findings. Impacts on PCS ED visits as a proportion of the impact on overall outpatient ED visits: Track 1: -3.5/-7.2 = 48 percent; Track 2: -4.7/-7.3 = 64 percent. More ED visits occur during the week, so impacts on weekday visits are larger in absolute terms. Impacts on weekday visits are also larger in relative terms (as a percentage of the comparison group mean) for Track 1.

[a] We report the actual, unadjusted averages in the baseline period which are similar for the CPC+ and comparison groups due to matching. In the intervention periods, the comparison group mean is computed by subtracting the regression-adjusted difference between the CPC+ and comparison means in each time period from the CPC+ mean in that same time period.

Table 5.L.3. ***(continued)***

[b] Each impact estimate is regression-adjusted using a difference-in-differences analysis that reflects the difference between the average outcome for Medicare FFS beneficiaries attributed to CPC+ practices in the first three years of CPC+ and the average outcome in the baseline year, relative to the same difference over time for Medicare FFS beneficiaries attributed to comparison practices, while controlling for beneficiary characteristics and practice fixed effects.

[c] We calculated percentage impacts relative to what the CPC+ mean would have been in PYs 1 through 3 (combined) in the absence of the intervention—that is, the unadjusted CPC+ mean minus the impact estimate.

[d] The sum of PCS outpatient ED visits is less than total outpatient ED visits because total outpatient ED visits include those for other care needs, such as injuries, mental health, drugs, and alcohol.

*/**/*** Significantly different from zero at the 0.10/0.05/0.01 level, two-tailed test.

C = comparison; ED = emergency department; FFS = fee-for-service; NA = not applicable; PCS = primary care substitutable; PY = Program Year; SE = standard error.

C. The effects of changes in access on changes in primary care substitutable ED visits for CPC+ practices

Changes in most components of practice-reported access were not associated with decreases in PCS ED visits. However, increases in one main component of access, same- or next-day appointment availability, explain a small portion of the Track 1 reductions in ED visits. For Track 1 practices, we estimate that improving from never offering same- or next-day appointments to offering them to the fullest extent to all patients (increasing their score from 0 to 1) would be associated with a 20.5 visit decrease in PCS ED visits per 1,000 beneficiaries ($p < 0.01$). Given that the average increase in the same- or next-day appointment score from 2017 to 2018 was 0.02 and CPC+ reduced PCS ED visit utilization relative to comparison practices by 3.5 visits per 1,000 beneficiaries among Track 1 practices, this suggests that this component of access only explains 10 percent of the PCS ED visit effect among Track 1 practices (Table 5.L.4).

Table 5.L.4. Associations between changes in access and changes in primary care substitutable ED visit rates

	Track 1			Track 2		
Measure of access	Coefficient (SE)	Mean change from PY 1 to PY 2	% of PCS ED effect explained[a]	Coefficient (SE)	Mean change from PY 1 to PY 2	% of PCS ED effect explained[a]
Overall access composite score	-10.0 (10.0)	0.04	12%	1.5 (9.6)	0.03	NA
Extended hours of operation	3.4 (3.8)	0.06	NA	1.1 (4.0)	0.03	NA
24/7 access to care team	-4.3 (6.7)	0.05	7%	0.8 (5.8)	0.03	NA
Same- or next-day appointments	-20.5*** (8.4)	0.02	10%	-1.0 (9.4)	0.01	0%
Percentage of primary care visits occurring on weekends	-1.5 (1.1)	0.10	4%	-2.1** (0.9)	0.03	2%

Source: Mathematica's analysis of Medicare claims data from January 2016 through December 2019; 2017 and 2018 CPC+ Practice Survey; and 2017 Q1–Q2 and 2018 Q3–Q4 CPC+ Practice Portal data. The CPC+ baseline period corresponds to 2016, Program Year (PY) 1 is 2017, PY 2 is 2018, and PY 3 is 2019. The CPC+ Practice Survey and CPC+ Practice Portal data were not available in 2016.

Note: Access composite scores could range from 0 to 1; all included variables were rescaled to 0 to 1 and then averaged to create the composite. Each row represents a separate regression model that included practice-level controls. Therefore, we generated these estimates from models that do not control for the other components of access, although they control for practice, geographic, and beneficiary characteristics at the practice level. Beneficiary demographic characteristic controls included average beneficiary age and the proportions for each of the following categories: age, race, sex, original entitlement reason, and dual eligibility. HCC scores and chronic condition controls included average HCC score and proportions for each of the following categories: Tier 4, Tier 5, diabetes, cancer, chronic obstructive pulmonary disease, chronic kidney disease, Alzheimer's and related dementia, heart failure, and behavioral health. Geographic characteristic controls included the median 2014 county-level household income, the 2015 HRR price index, an indicator for a primary care health professionals shortage area, the quartiles of total hospital beds in 2013 per 10,000 2014 county-level total population, the proportions of rural and suburban statuses, the percentage of county-level poverty in 2014, the percentage of county-level Medicare Advantage enrollees/eligible beneficiaries in 2015, and the percentage of people age 25+ with 4 years of college (2010–2014). The practice characteristic controls included an indicator for ownership by a health system or hospital, categorical counts of primary care practitioners, categorical counts of providers, an indicator for participation in the Shared Savings Program as of January 1, 2017, an indicator for experience with a prior transformation (e.g., participated in MAPCP, medical home recognition, or participated in CPC Classic), an indicator for whether the practice has providers from multiple specialties, and the percentage of charges that are primary care.

[a] We calculated the percentage of effect explained by taking [(estimated impact*mean change from PY 1 to PY 2)/PCS ED effect coefficient]*100. For Track 1, the PCS ED effect coefficient was -3.5, for Track 2, the PCS ED effect coefficient was -4.7. In some cases, the estimated impact and mean change were both positive. Therefore, the resulting percentage of the PCS ED effect explained could be negative given the PCS ED effect denominator is negative for both tracks; in those cases, the percentage of the PCS ED effect explained is shown as NA.

*/**/*** Significantly different from zero at the 0.10/0.05/0.01 level, two-tailed test.

ED = emergency department; HCC = hierarchical condition category; HRR = hospital referral region; MAPCP = Multi-payer Advanced Primary Care Practice; NA = not applicable; PCS = primary care substitutable; PY = Program Year; SE = standard error.

Care manager staffing for Track 2 CPC+ practices also explained a small portion (11 percent) of the PCS ED visit effect. Increases in Track 2 care manager staffing FTEs per 1,000 patients were associated with 10.4 fewer PCS ED visits per 1,000 beneficiaries ($p < 0.01$, Table 5.L.5). We did not find evidence that changes in PCP mid-level staffing were associated with reductions in ED visits, but we did find evidence that Track 2 practices that had larger increases in PCP physician staffing had larger decreases in PCS ED visits. However, since the average change in PCP physician staffing across all CPC+ practices was negative, changes in PCP physician staffing do not explain the 4.7 visit reduction effect we found in PCS ED visits, even though increasing PCP physician staffing appears to be a factor in reducing PCS ED visits.[130]

Table 5.L.5. Associations between changes in staffing and changes in primary care substitutable ED visit rates

	Track 1			Track 2		
Supplemental measure of access	Coefficient (SE)	Mean change from PY 1 to PY 2	% of PCS ED effect explained[a]	Coefficient (SE)	Mean change from PY 1 to PY 2	% of PCS ED effect explained[a]
Number of PCP physician FTEs per 1,000 patients	-0.6 (1.0)	-0.13	NA	-2.7* (1.5)	-0.07	NA
Number of PCP mid-level FTEs per 1,000 patients	3.1 (3.0)	0.01	NA	-3.7 (2.7)	0.03	2%
Number of care manager FTEs per 1,000 patients	-0.1 (2.1)	0.09	0%	-10.4*** (2.5)	0.05	11%

Source: Mathematica's analysis of Medicare claims data from January 2016 through December 2019; 2017 and 2018 CPC+ Practice Survey; and 2017 Q1–Q2 and 2018 Q3–Q4 CPC+ Practice Portal data. The CPC+ baseline period corresponds to 2016, Program Year (PY) 1 is 2017, PY 2 is 2018, and PY 3 is 2019. The CPC+ Practice Survey and CPC+ Practice Portal data were not available in 2016.

Note: Access composite scores could range from 0 to 1; all included variables were rescaled to 0 to 1 and then averaged to create the composite. Each row represents a separate regression model that included practice-level controls. Therefore, we generated these estimates from models that do not control for the other components of access, although they control for practice, geographic, and beneficiary characteristics at the practice level. Beneficiary demographic characteristic controls included average beneficiary age and the proportions for each of the following categories: age, race, sex, original entitlement reason, and dual eligibility. HCC scores and chronic condition controls included average HCC score and proportions for each of the following categories: Tier 4, Tier 5, diabetes, cancer, chronic obstructive pulmonary disease, chronic kidney disease, Alzheimer's and related dementia, heart failure, and behavioral health. Geographic characteristic controls included the median 2014 county-level household income, the 2015 HRR price index, an indicator for a primary care health professionals shortage area, the quartiles of total hospital beds in 2013 per 10,000 2014 county-level total population, the proportions of rural and suburban statuses, the percentage of county-level poverty in 2014, the percentage of county-level Medicare Advantage enrollees/eligible beneficiaries in 2015, and the percentage of people age 25+ with 4 years of college (2010–2014). The practice characteristic controls included an indicator for ownership by a health system or hospital, categorical counts of primary care practitioners, categorical counts of providers, an indicator for participation in the Shared Savings Program as of January 1, 2017, an indicator for experience with a prior transformation (e.g., participated in MAPCP, medical home recognition, or participated in CPC Classic), an indicator for whether the practice has providers from multiple specialties, and the percentage of charges that are primary care.

[a] We calculated the percentage of effect explained by taking [(estimated impact*mean change from PY 1 to PY 2)/PCS ED effect coefficient]*100. For Track 1, the PCS ED effect coefficient was -3.5, for Track 2, the PCS ED effect coefficient was -4.7. In some cases, the estimated impact and mean change were both positive or both negative. Therefore, the resulting percentage of the PCS ED effect explained could be negative given the PCS ED effect denominator is negative for both tracks; in those cases, the percentage of the PCS ED effect explained is shown as NA.

*/**/*** Significantly different from zero at the 0.10/0.05/0.01 level, two-tailed test.

ED = emergency department; FTE = full-time equivalent; HCC = hierarchical condition category; HRR = hospital referral region; MAPCP = Multi-payer Advanced Primary Care Practice; NA = not applicable; PCP = primary care practitioner; PCS = primary care substitutable; PY = Program Year; SE = standard error.

[130] This suggests that for average practices, decreases in PCP physician staffing FTEs per 1,000 patients were associated with increased PCS ED visits per 1,000 beneficiaries.

Increases in the percentage of primary care visits occurring on weekends are associated with reductions in Track 2 PCS ED visits on weekdays and non-weekdays alike. For Track 2 CPC+ practices, increases in the percentage of PC visits occurring on weekends, a proxy for weekend access to the practice, were associated with 1.7 fewer weekday PCS ED visits ($p < 0.05$, Table 5.L.6) and 0.7 fewer non-weekday PCS ED visits ($p < 0.05$, Table 5.L.6) per 1,000 beneficiaries. The larger reduction in absolute terms on weekdays of PCS ED visits associated with this practice weekend availability proxy is counterintuitive and may suggest that other practice changes that are correlated with extending weekend hours might be driving these results.

Table 5.L.6. Association between the proxy for weekend availability and weekday and non-weekday primary care substitutable ED visit rates

	Track 1			Track 2		
Supplemental measure of access	Coefficient (SE)	Mean change from PY 1 to PY 2	% of PCS ED effect explained[a]	Coefficient (SE)	Mean change from PY 1 to PY 2	% of PCS ED effect explained[a]
Outcome: Change in weekday PCS ED visits						
Change in percentage of primary care visits on weekend	-1.2 (0.8)	0.10	4% of -3.1 weekday PCS ED visits	-1.7** (0.7)	0.03	2% of -3.1 weekday PCS ED visits
Outcome: Change in non-weekday (weekends and holidays) PCS ED visits						
Change in percentage of primary care visits on weekend	-0.7 (0.4)	0.10	15% of -0.5 weekend PCS ED visits	-0.7** (0.4)	0.03	2% of -1.5 weekend PCS ED visits

Source: Mathematica's analysis of Medicare claims data from January 2016 through December 2019. The CPC+ baseline period corresponds to 2016, Program Year (PY) 1 is 2017, PY 2 is 2018, and PY 3 is 2019.

Note: Access composite scores could range from 0 to 1; all included variables were rescaled to 0 to 1 and then averaged to create the composite. We generated these estimates from models that do not control for the other components of access, although they control for practice, geographic, and beneficiary characteristics at the practice level. Beneficiary demographic characteristic controls included average beneficiary age and the proportions for each of the following categories: age, race, sex, original entitlement reason, and dual eligibility. HCC scores and chronic condition controls included average HCC score and proportions for each of the following categories: Tier 4, Tier 5, diabetes, cancer, chronic obstructive pulmonary disease, chronic kidney disease, Alzheimer's and related dementia, heart failure, and behavioral health. Geographic characteristic controls included the median 2014 county-level household income, the 2015 HRR price index, an indicator for a primary care health professionals shortage area, the quartiles of total hospital beds in 2013 per 10,000 2014 county-level total population, the proportions of rural and suburban statuses, the percentage of county-level poverty in 2014, the percentage of county-level Medicare Advantage enrollees/eligible beneficiaries in 2015, and the percentage of people age 25+ with 4 years of college (2010–2014). The practice characteristic controls included an indicator for ownership by a health system or hospital, categorical counts of primary care practitioners, categorical counts of providers, an indicator for participation in the Shared Savings Program as of January 1, 2017, an indicator for experience with a prior transformation (e.g., participated in MAPCP, medical home recognition, or participated in CPC Classic), an indicator for whether the practice has providers from multiple specialties, and the percentage of charges that are primary care.

[a] We calculated the percentage of effect explained by taking [(estimated impact*mean change from PY 1 to PY 2 1)/PCS ED effect coefficient]*100.

*/**/*** Significantly different from zero at the 0.10/0.05/0.01 level, two-tailed test.

ED = emergency department; HCC = hierarchical condition category; HRR = hospital referral region; MAPCP = Multi-payer Advanced Primary Care Practice; NA = not applicable; PCS = primary care substitutable; PY = Program Year; SE = standard error.

5.L.4. Discussion

We found that access improved over the first three years of the model and that there were reductions in PCS ED visit utilization for CPC+ practices compared to comparison practices. However, the changes in access that we were able to measure only explained a small fraction of the improvements in PCS ED visit utilization. The results for the associations between different measures of access and PCS ED visits were also inconsistent across tracks. For example, increases in same- or next-day appointment availability were significantly associated with decreased PCS ED visit use for Track 1 only, and increases in care manager staffing FTEs per 1,000 patients and the percentage of PC visits occurring on weekends were significantly associated with decreased PCS ED visit use for Track 2 only. In addition, increases in the percentage of PC visits occurring on the weekends had larger associations with changes in weekday PCS ED visits than weekend visits. One explanation for the inconsistent associations we found is that the measures of access could be co-occurring with other improvements in care delivery that we did not control for. For example, the CPC+ Study of Exemplar Practices identified some characteristics that facilitated exemplar practices' ability to implement acute hospitalization utilization reduction strategies, such as organizational support for and staff interest in innovation (Laird et al. 2022, Appendix 4.C). We did not control for such support and innovation that might also impact the need for outpatient ED visits. Furthermore, the secondary measure of access of care manager staffing could be measuring aspects of care delivery other than access (such as management of chronic conditions preventing the need for outcome utilization rather than access to care for those conditions).

There are some additional limitations with our analysis. There could be errors in practice reporting for the CDR Portal and CPC+ Practice Survey data sources we used to measure access, and it is possible that practices might have made early changes related to the model before they answered the first CPC+ Practice Survey. These issues could both lead to our estimates understating the true association of access with PCS ED visits. In addition, given the limited scope of the CDR Portal and CPC+ Practice Survey questions, there could also be aspects of access that we did not capture, such as unscheduled alternatives to in-person office visits (e.g., same-day phone or video visit availability for an urgent need) that are important to reducing PCS ED visits. Lastly, we controlled for many practice-level characteristics in our regression models, but there could still be factors we cannot measure that contribute to practice's decreases in ED visits and are related to their decisions to improve access, which would lead us to overestimate the impact of improvements in access on decreases in ED visits.

5.L.5. Conclusion

Our results do not provide clear evidence that improvements in access led to improvements in ED visits, and in fact, we find that improvements in access—in the way we can measure—explain very little of the effect of CPC+ on ED visit utilization. To better understand the mechanisms through which changes in access affects outcomes, researchers and model evaluators need better measures of access available in all time periods included in the study that are measured and answered consistently over time. In addition to better measures of access, if policymakers want to be able to understand which care processes lead to improved outcomes, future innovation models should explicitly test different care delivery process changes.

5.M. Michigan quantitative analysis

Key takeaway

The "CPC+ Michigan ED and Inpatient Utilization High-Performing Practice Study" qualitatively identified six practice transformation themes among Michigan CPC+ practices that were high-performers based on low emergency department (ED) and inpatient utilization. We used quantitative methods among all CPC+ practices and found that improvements in some, but not all, of the Michigan qualitative study's practice transformation themes were associated with decreased outpatient ED visit and acute hospitalization utilization. Improvements in practices' availability and responsiveness to patient needs (i.e., offering same- or next-day appointments and after-hours access) and identification of patients needing intervention (i.e., using a standard, integrated method to stratify patients by risk level and registry data to identify and manage groups of patients) showed the strongest, most robust effects on improving these outcomes. The consistency of the results for these themes using both the Michigan study's qualitative evidence and our quantitative analyses' evidence suggests that it is important to invest resources in these care delivery changes.

5.M.1. Introduction

Jerome Finkel, MD and Diane Marriott, DrPH produced a qualitative study[131] for the CPC+ Michigan Multistakeholder Care Interventions Subcommittee. They used Blue Cross Blue Shield of Michigan commercial data and CMS Medicare FFS beneficiary claims data[132] to identify 10 practices that they deemed to be high performers because they had the lowest 20 percent unadjusted[133] utilization for both ED visits and inpatient general hospital/acute care utilization (inpatient visit utilization) among Michigan CPC+ practices from October 1, 2017 to September 30, 2018. They also conducted a literature review and gathered expert input to identify national high-performing health systems. They interviewed the Michigan practices[134] as well as the high-performing health systems and asked them what techniques, structures, and processes of care they perceived to have contributed to their lower ED and inpatient visit utilization. They found six practice characteristics or interventions (themes) that the practices and health systems attributed to having better ED and inpatient visit utilization outcomes:[135]

[131] In this summary document, we referred to this study hereafter as "the Michigan qualitative study."

[132] Authors noted leveraging the multipayer CPC+ Michigan dashboard (CMS, Blue Cross Blue Shield, and Priority Health) to identify high-performing Michigan CPC+ practices in the blog post, but the linked summary findings document noted that "the claims data available at the time of analysis was limited to BCBSM commercial and CMS patients," so it is unclear whether Priority Health claims were included (Finkel and Marriott, 2021).

[133] Authors noted a limitation that risk adjustment was not possible with the study's data source, but that "proxies for differences in underlying patient risk burdens were used including comparison of mean patient age and Medicare HCC score" (Finkel and Marriott 2021).

[134] For the interviews, researchers asked practices to involve staff in a variety of roles, including at least one clinical partner and one administrative partner who were familiar with the practice and its processes.

[135] See Finkel and Marriott (2021). For more details, see Table 5.M.1.

- **Theme 1: Physician and medical director/clinician lead engagement in CPC+ culture and innovations.** "Physician engagement drives patient and practice team engagement and promotes a practice culture that embraces adapting innovations to improve care…"
- **Theme 2: Co-located teams with care management.** "Co-located, engaged teams with care management at the core are key."
- **Theme 3: Offloading routine tasks.** "Offloading routine tasks (e.g., medication refills, gap closures) from the PCP workstream frees physicians to focus on patient needs and championing team-based care."
- **Theme 4: Availability and responsiveness to patient needs** (and patient awareness of availability). "Availability and responsiveness to patient needs as well as patient awareness of the availability mattered more than extended hours…[This included] the patient's ability to have clinical expertise that responds to patient questions and needs quickly. Patient calls were returned the same day, and in some high-performers, within the hour."
- **Theme 5: Integrated and regular performance reporting.** "Performance reporting integrated in regular team huddles or communication drives attention to and accountability for performance. Sharing provider-level performance regularly similarly motivated improvement among individual providers."
- **Theme 6: Identification of patients needing intervention.** "High performing practices had a method for identifying patients that would benefit from interventions (e.g., care management, self-management programs; remote patient monitoring; etc.). All high-performers studied readily recited their "triggers" for intervention and care management."

While the Michigan qualitative study presents valuable evidence on the potential of these themes for improving primary care, there are several limitations to its methodological approach, which warranted further investigation. First, the practices likely did not have empirical evidence of the causal mechanisms and pathways leading to lower utilization, so they might have attributed their beneficiaries' utilization to certain activities when other factors could have been responsible. In addition, even if the cited activities helped practices to lower utilization, the activities may only work in their particular setting (i.e., Michigan CPC+ practices or health systems). Therefore, it was unclear whether the Michigan qualitative study's themes could be generalized to all CPC+ practices. The motivation for our analyses was to use these promising mechanisms of action to reduce acute care utilization and determine whether they generalize to all CPC+ practices. We quantitively tested these hypothesized mechanisms of action using a rigorous methodological approach that allowed for controlling for other observable practice characteristics.[136]

[136] The Michigan study had a similar methodology to the one used in the CPC+ exemplar study; however, to identify exemplars, our team identified the practices with the highest probability of achieving substantial reductions in risk-adjusted Medicare acute hospitalization rate (AHR) between baseline and PY 2 using Bayesian methods. We also plan to do similar quantitative analyses to see if the exemplar findings can be generalized to the full CPC+ sample with a quantitative approach.

5.M.2. Data, sample, and methods

We first identified proxies for the six themes using CPC+ Practice Survey data to create composite theme scores ranging from 0 to 1 (see Table 5.M.1).[137] Figure 5.M.1 shows trends over time in the theme scores among CPC+ 2017 Starter Track 1 and 2 practices.

We then assessed whether the 10 high-performing CPC+ practices included in the Michigan qualitative study had higher levels of these theme scores in Program Year (PY) 2 (2018) compared to other CPC+ practices nationwide.[138] We also used Medicare claims data to identify high-performing practices that had the lowest 20 percent unadjusted utilization for both outpatient ED visits and acute hospitalizations during PY 2 among the full nationwide sample of CPC+ practices.[139] We similarly compared theme scores for this group of high performers relative to all other CPC+ practices.

Then, using the full sample of CPC+ practices, for the first three program years (2017–2019), we used cross-sectional regressions to examine whether practices with higher theme scores in that particular year had lower outpatient ED visit and acute hospitalization utilization in that year, while controlling for baseline beneficiary, practice, and geographic factors.[140,141]

[137] For each variable included in the proxies for the theme, we recoded the variable to be a binary (0 or 1) top-box or rescale to be "pseudo-continuous" (between 0 and 1). For measures that included more than one variable, we then calculated composite scores by taking the mean across all survey items included in the activity for each practice.

[138] The Michigan qualitative study listed the name, address, and physician organization for the practices included in the study's interviews. Mathematica used CPC+ portal information to cross-walk those practices to their CPC+ practice IDs to identify them in our analyses.

[139] We used unadjusted outcomes to align with the Michigan qualitative study's approach to identifying high performing practices using unadjusted outcomes.

[140] Results were similar in sensitivity models that did not control for any baseline practice characteristics, such as those that could drive transformations captured by the themes we investigated in this analysis (e.g., SSP participation, hospital or system ownership, etc.). Results were sensitive to excluding all control variables, however.

[141] Due to data constraints at the time this analysis was conducted, we only analyzed data through PY 3. As a follow-up, Appendix 5.O. Synthesis describes results from regressions of changes in promising care processes from PY 1 to PY 2 on changes in outcomes from PY 2 through PY 5.

Table 5.M.1. Proxy source variables for theme composite scores

Theme	Michigan qualitative study theme description[a]	Mathematica's quantitative study CPC+ Practice Survey proxy source variables[b]	Potential limitations of proxy variables
Theme 1: Physician/medical director/clinician lead engagement in CPC+ culture and innovations	"Physician engagement drives patient and practice team engagement and promotes a practice culture that embraces adapting innovations to improve care regardless of setting (large or small practice; part of a health system or independent)."	1. **Medical director/clinician lead involvement in CPC+,** recoded top-box variable: Thinking of the different types of staff at this practice site, how involved is each staff type in implementing CPC+? Medical director or clinician lead at this practice site is…0=somewhat involved, not very involved, or not at all involved. 1=very involved. 2. **Physician involvement in CPC+,** rescaled pseudo-continuous variable: Thinking of the different types of staff at this practice site, how involved is each staff type in implementing CPC+? Physicians: 0=not at all involved, 0.33=not very involved, 0.66=somewhat involved. 1=very involved.	These proxy variables might not capture the innovation aspect (beyond implementing CPC+) of this Michigan qualitative study theme.
Theme 2: Co-located teams with care management	"Co-located, engaged teams with care management at the core are key. The size of team does not matter but co-location does."	1. **Practice's care managers provide services on location for high-risk patients,** recoded top-box variable: Care management services for high-risk patients…0=are not provided at this practice, or are provided by care managers from an outside organization (e.g., a health insurance plan), or are provided by a care manager within this practice's organization who is not physically located at this practice site. 1=are provided by a care manager located at this practice site. 2. **Extent to which practice site's care team huddles happen,** rescaled pseudo-continuous variable: At this practice site, care team huddles…0=do not occur, 0.33=occur some days, 0.66=occur most days, 1=occur every day.	The team engagement with care management aspect of this theme might not be perfectly measured by these survey variables.

Table 5.M.1. ***(continued)***

Theme	Michigan qualitative study theme description[a]	Mathematica's quantitative study CPC+ Practice Survey proxy source variables[b]	Potential limitations of proxy variables
Theme 3: Offloading routine tasks	"Offloading routine tasks (e.g., medication refills, gap closures) from the PCP workstream frees physicians to focus on patient needs and championing team-based care."	1. **Non-physicians' performance of key clinical service roles,** recoded top-box variable: Non-physician practice team members…0=play a limited role in providing clinical care, or are primarily tasked with managing patient flow and triage, or provide some clinical services such as assessment or self-management support. 1=perform key clinical service roles that match their abilities and credentials. 2. **Extent to which pre-visit planning is done,** rescaled pseudo-continuous variable: Pre-visit planning (gathering and organizing patient information to prepare for the visit) prior to the day of the visit…0=is not done; 0.33=is done but primarily focuses on reviewing test results and consultation reports from specialist referrals; 0.66=is done and includes (1) reviewing test results and consultation reports from specialist referrals, and (2) identifying gaps in health care (e.g., a needed flu shot or cancer screenings); 1=is done and includes (1) reviewing test results and consultation reports from specialists, (2) identifying gaps in health care, and (3) conducting outreach before the visit, to ask the patient to obtain needed tests prior to the visit.	The second proxy variable assumes pre-visit planning allows physicians the extra time to focus on patient needs and team-based care.
Theme 4: Availability and responsiveness to patient needs (and patient awareness of availability)	"Availability and responsiveness to patient needs as well as patient awareness of the availability mattered more than extended hours. Though hours outside traditional 8am-5pm practice operations can be very helpful for those whose schedules cannot accommodate standard workweek hours, they are not useful to patients if they are consistently filled or cannot accommodate an urgent need. More important is the patient's ability to have clinical expertise that responds to patient questions and needs quickly. Patient calls were returned the same day, and in some high-performers, within the hour."	1. **Same-day appointment availability,** recoded top-box variable: Same-day appointments for patients who need them are available at this practice site for…0=none of this practice's patients, or some of this practice's patients, or many of this practice's patients. 1=most or all of this practice's patients. 2. **After-hours access,** recoded top-box variable: Patient after-hours access (24 hours, 7 days a week) to a physician, PA/NP, or nurse…0=is not available or is limited to an answering machine; or is available from a coverage arrangement (e.g., answering service) that does not offer a standardized communication protocol back to the practice for urgent problems; or is provided by a coverage arrangement (e.g., answering service) that shares necessary patient data with and provides a summary to the practice. 1=is available via the patient's choice of email or phone directly with the practice team or a practitioner who has real-time access to the patient's electronic medical record.	It is possible these proxy variables do not fully capture patient awareness of the availability described in the Michigan qualitative study theme, and these CPC+ Practice Survey questions don't capture how quickly practices respond to patient needs.

Table 5.M.1. ***(continued)***

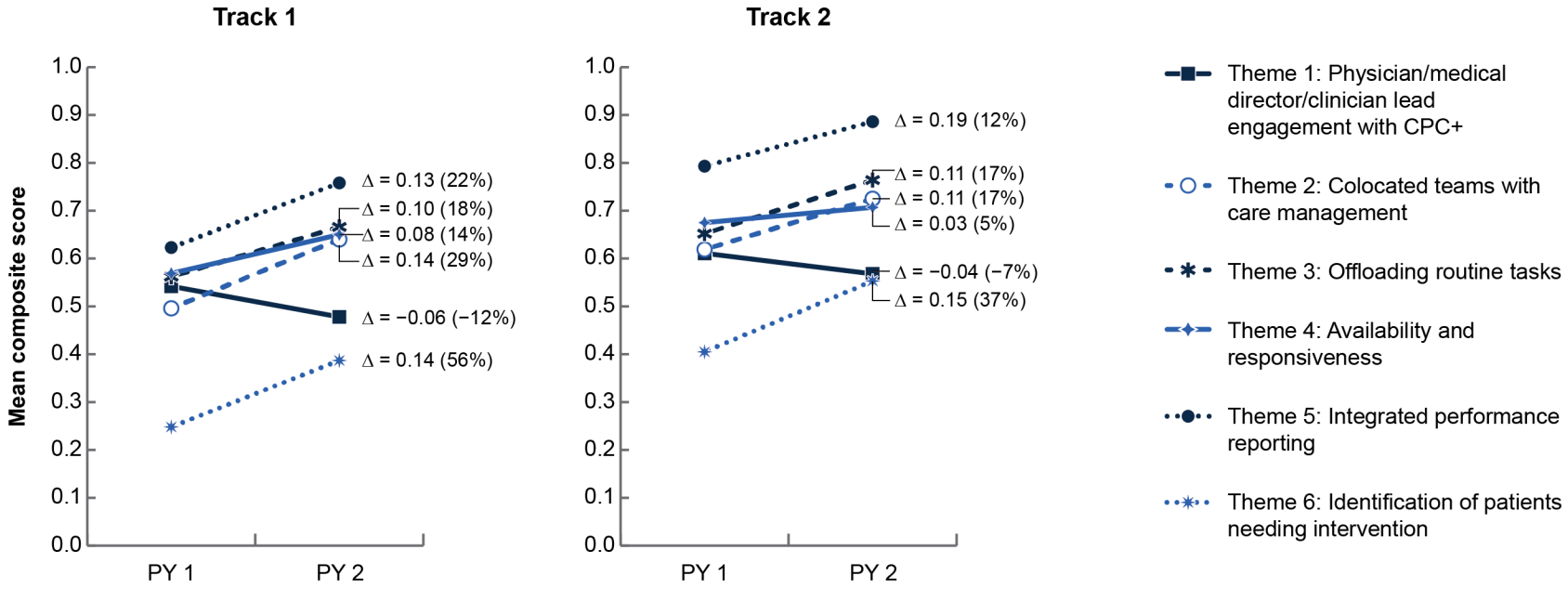
Theme	Michigan qualitative study theme description[a]	Mathematica's quantitative study CPC+ Practice Survey proxy source variables[b]	Potential limitations of proxy variables
Theme 5: Integrated and regular performance reporting	"Performance reporting integrated in regular team huddles or communication drives attention to and accountability for performance. Sharing provider-level performance regularly similarly motivated improvement among individual providers."	1. **Use of performance measures to guide quality improvement,** recoded top-box variable: Use of performance measures by this practice site to guide quality improvement (QI)…0=is not done, or is rarely done, or is sometimes done. 1=is usually done.	This theme is composed of only one proxy source variable; this CPC+ Practice Survey question is high level enough that it could capture provider-level performance, but it isn't specific to the provider level.
Theme 6: Identification of patients needing intervention	"High performing practices had a method for identifying patients that would benefit from interventions (e.g., care management, self-management programs; remote patient monitoring; etc.) All high-performers studied readily recited their "triggers" for intervention and care management."	1. **Use of patient stratification by risk level,** recoded top-box variable: A standard method or tool(s) to stratify patients by risk level…0=is not available, or is available but not consistently used to stratify all patients, or is available and is consistently used to stratify all patients but is inconsistently integrated into all aspects of care delivery. 1=is available, consistently used to stratify all patients, and is integrated into all aspects of care delivery. 2. **Use of registry data to assess or manage groups of patients,** recoded top-box variable: A registry is a data system that identifies and tracks patients with specific health conditions, risk states, or medications. At this practice site, registry data to assess or manage care for groups of patients…0=are not available, or are available for 1–2 diseases and/or risk states, or are available for 3– 5 diseases and/or risk states. 1=are available for 6 or more diseases and/or risk states.	These questions capture the theme well, but they do assume that the interventions or additionally needed care actually occurs.

Source: Finkel and Marriott (2021).

[a] Mathematica's analysis of the 2017, 2018, and 2019 CPC+ Practice Survey.

[b] All questions are sourced from the CPC+ Practice Survey. Theme 5 data are not available in Program Year 3 (2019).

Figure 5.M.1. Trends over time in theme scores among CPC+ 2017 Starter Track 1 and 2 practices



Source: Mathematica's analysis of the 2017 and 2018 CPC+ Practice Survey.

Note: Theme scores (and their means) ranged from 0 to 1, with 1 being the highest score possible and indicating the practices used the process or offered the service measured by the theme's CPC+ Practice Survey proxy variable(s) to the fullest extent possible. The CPC+ baseline period corresponds to 2016, Program Year (PY) 1 is 2017, PY 2 is 2018, and PY 3 is 2019. The CPC+ Practice Survey data were not available in 2016. Additionally, data we used to measure Theme 5 (Integrated performance reporting) were not available in PY 3.

PY = Program Year.

The cross-sectional results could still be biased by practice characteristics that we cannot measure but are related to the practices having higher theme scores and lower utilization. If a theme had a causal effect, we would expect improvement in the theme to be associated with improvements in the outcome. Therefore, we also used a more rigorous "changes-in-changes" ordinary least squares regression approach. However, there is more limited variation in changes in themes compared to the theme scores used in the cross-sectional model, and changes in themes may take more time to result in changes in utilization than we were able to measure with our data sources.[142] Therefore, these models tested whether practices with *greater improvements* in these themes from PY 1 (2017, the earliest available CPC+ Practice Survey data) to PY 2 also had larger reductions in their outpatient ED visit and acute hospitalization utilization from baseline (2016) to PY 3 (2019).[143] These changes-in-changes models controlled for the baseline theme score,[144] utilization rate, and beneficiary, practice, and geographic factors.[145]

For both the cross-sectional and changes-in-changes models, we included only 2017 Starter CPC+ practices, separately by track. We estimated separate regressions for each theme and utilization outcome.[146] These models all used weights accounting for the practice's baseline beneficiary count and clustered standard errors at the practice level. Table 5.M.2 summarizes the data sources and years we included in each of our analyses.

[142] Using CPC+ Practice Survey data to measure themes, we observed increases over time in both tracks for all themes except for physician and medical director/clinician lead engagement in CPC+ culture and innovations (Theme 1, Figure B.1).

[143] We examined changes in themes from PY 1 to PY 2 (2017 to 2018) to allow time for the practices' care delivery changes to impact the beneficiary service utilization outcomes through PY 3, and to remove the possibility of reverse causality (i.e., ED visit outcomes in PY 3 informing practices' determination of practice transformation themes in PY 3).

[144] There was one exception: the Theme 5 (integrated and regular performance reporting) models did not control for the baseline theme score due to collinearity concerns because that theme score was composed of only one survey question.

[145] We ran models to assess sensitivity to the set of control variables we used and generally found similar results. Starting with unadjusted results only controlling for the baseline theme score and baseline outcome rate, we progressively added more control variables by category (i.e., baseline beneficiary demographics, baseline hierarchical condition category scores and chronic conditions, baseline geographic characteristics, baseline practice characteristics [main model results], and change in beneficiary characteristics [not included in main model results]).

[146] We also ran sensitivity models that controlled for the different themes in the same model and generally found similar results. The Michigan qualitative study identified practices with lower utilization for both ED and inpatient care utilization, whereas we had separate regressions for each of the outpatient ED visit utilization rate and the acute hospitalization utilization rate outcomes.

Table 5.M.2. Sample, data sources, and years by analysis

Analysis	Data sources and years
High-performing practices' theme scores	• Theme scores: CPC+ Practice Survey: 2018[a] • Identifying high-performing practices: – Michigan qualitative study's list of included practices – Mathematica's analysis of Medicare claims from January 2018 through December 2018[b]
Cross-sectional	• Outcomes: ED visit utilization and inpatient utilization rates based on Medicare claims from January 2017 through December 2019[a] • Key explanatory variables: theme scores based on CPC+ Practice Survey from 2017, 2018, and 2019
Changes-in-changes	• Outcomes: Changes in ED utilization and changes in inpatient utilization from 2016 to 2019, using Medicare claims from January 2016 through December 2019[a,d] • Key explanatory variables (themes): Changes in theme scores from 2017 to 2018, using CPC+ Practice Survey: 2017 and 2018 [a,c]

[a] The CPC+ baseline period corresponds to 2016, Program Year (PY) 1 is 2017, PY 2 is 2018, and PY 3 is 2019.

[b] The Michigan qualitative study identified high-performing Michigan CPC+ practices that had the lowest 20 percent unadjusted utilization for both outpatient ED visit and inpatient general hospital/acute care utilization from 10/1/2017 to 9/30/2018. We assessed mean theme scores during PY 2 among the 10 practices identified in the Michigan qualitative study compared to all other CPC+ practices nationwide. We also assessed mean theme scores during PY 2 among the CPC+ practices Mathematica identified as having the lowest 20 percent outpatient ED visit and acute hospitalization utilization nationwide in PY 2 among the full sample of CPC+ practices.

[c] We examined changes in practice transformation themes from PY 1 to PY 2 to allow time for the practices' care delivery changes to impact the beneficiary service utilization outcomes through PY 3 and to remove the possibility of reverse causality (i.e., ED visit outcomes in PY 3 informing practices' determination of practice transformation themes in PY 3).

ED = emergency department; PY = Program Year.

5.M.3. Results

A. High-performing practices' theme scores

Among the 10 high-performing practices included in the Michigan qualitative study, our analysis of the PY 2 CPC+ Practice Survey data showed that they had higher scores for four of the six themes compared to all other CPC+ practices.[147] Findings from this independent data source verify one aspect of the Michigan qualitative study—that these practices were exceptional in terms of most of these themes. The themes for which these Michigan practices had lower scores compared to other CPC+ practices were Theme 1 (physician/medical director/clinician lead engagement in CPC+ culture and innovations) and Theme 3 (offloading routine tasks, Table 5.M.3). Theme 1 scores decreased from PY 1 to PY 2, which may be because most engagement among physicians, medical directors, and clinician leads occurred in the planning stage, and perhaps these individuals did not need to be as engaged in later years. For Theme 3, one of the two proxy variables we used assumes pre-visit planning allows physicians the extra time to focus on patient needs and team-based care, so the lower scores may reflect that our measure did not fully capture the theme.

[147] We used PY 2 (2018) data in this descriptive comparison because the Michigan qualitative study identified high-performing Michigan practices that had the lowest 20 percent unadjusted utilization for both ED and inpatient visits from 10/1/2017 to 9/30/2018.

Among the group of 217 CPC+ practices that we identified as high performers in the full nationwide sample of CPC+ practices, only two of the themes had higher scores compared to all other CPC+ practices (Theme 4: availability and responsiveness, and Theme 5: integrated performance reporting). In addition, the differences for those two themes between these 217 CPC+ high performers and all other CPC+ practices were smaller and less meaningful than the differences were between the Michigan qualitative study practices and all other practices (Table 5.M.3). Given the differences were smaller and these high performers had higher scores for fewer themes, this might suggest that the themes are particular to specific practices included in the Michigan qualitative study and less generalizable mechanisms for reducing utilization.

Table 5.M.3. Mean theme scores during Program Year 2, among CPC+ 2017 Starter Track 1 and 2 practices: Michigan qualitative study practices and nationwide CPC+ "high-performer" practices compared to all other CPC+ practices

Theme composite mean score in 2018[a]	CPC+ practices in Michigan qualitative study (n = 10)	All other CPC+ practices nationwide (n varies by theme, ranges from 2,741–2,755)	Nationwide high-performer practices with lowest 20% ED and IP utilization (n = 217)	Non-high-performer CPC+ practices nationwide (n varies by theme, ranges from 2,534–2,548)
Theme 1: Physician and clinician lead/medical director engagement with CPC+	0.42	0.51	0.50	0.51
Theme 2: Co-located teams with care management	0.72	0.71	0.70	0.71
Theme 3: Offloading routine tasks	0.63	0.73	0.71	0.73
Theme 4: Availability and responsiveness	0.88	0.69	0.71	0.69
Theme 5: Integrated performance reporting	0.95	0.85	0.86	0.85
Theme 6: Identification of patients needing intervention	0.70	0.49	0.45	0.49

Source: Mathematica's analysis of the 2018 CPC+ Practice Survey.

Note: The CPC+ Program Year (PY) 2 is 2018.The Michigan qualitative study identified high-performing Michigan CPC+ practices that had the lowest 20 percent unadjusted utilization for both outpatient ED visit and inpatient general hospital/acute care utilization from October 1, 2017, to September 30, 2018. Therefore, this table shows mean theme scores during PY 2 among the 10 practices identified in the Michigan qualitative study compared to all other CPC+ practices nationwide. This table also shows the means during PY 2 among the CPC+ practices Mathematica identified as having the lowest 20 percent outpatient ED visit and acute hospitalization utilization nationwide in PY 2 among the full sample of CPC+ practices.

[a] Theme scores (and their means) ranged from 0 to 1, with 1 being the highest score possible and indicating the practices used the process or offered the service measured by the theme's CPC+ Practice Survey proxy variable(s) to the fullest extent possible.

ED = emergency department; IP = inpatient; PY = Program Year.

B. Associations of high-performing practice themes with ED and inpatient utilization

Findings from our cross-sectional models did not uniformly substantiate the Michigan qualitative study's findings, although there were some consistent favorable effects primarily concentrated among Track 1 practices' ED utilization. For all themes, we found a significant negative relationship for at least one outcome (in the expected direction where higher composite scores were associated with reduced service use), but we didn't find complete consistency for any theme across outcomes or tracks. Most significant negative associations with the themes were primarily among Track 1 practices' outpatient ED visit utilization. For example, the maximum Theme 1 score of 1 compared to the minimum score of 0 (i.e., a practice whose physician and clinician lead/medical director are engaged with CPC+ to the highest extent compared to a practice whose physician and clinician lead/medical director are not engaged at all), is associated with 13.1 fewer ED visits per 1,000 beneficiaries among CPC+ Track 1 practices (p = 0.01, Table 5.M.4). Although Themes 3 (offloading routine tasks) and 5 (integrated performance reporting) for Track 2 ED utilization showed significant positive associations between *higher* theme scores and *higher* utilization, most significant findings supported our hypothesis that higher theme scores were associated with lower utilization (Table 5.M.4).[148] There could still be bias in these results due to any factors that could be related to practices' theme scores and utilization that we could not capture with our practice controls. The amount of bias in the cross-sectional model was unclear, which supported the importance of also examining a changes-in-changes model.

[148] We estimated separate regressions for each theme and utilization outcome. However, we also ran sensitivity models that controlled for the different themes in the same model and tested whether they were jointly significant (i.e., using a null hypothesis that the estimates for all themes were zero). For the ED visit utilization outcome, the themes were jointly significant at the 5% level for Track 1 practices, but not jointly significant for Track 2 practices. For the acute hospitalization outcome, the themes were jointly significant at the 5% level for Track 1 and 1% level for Track 2 practices.

Table 5.M.4. Cross-sectional association between themes and outcomes (outpatient ED visit or acute hospitalization rate) during the first three Program Years among CPC+ practices by track

	Mean (SD) Theme score, PY 1 to 3		Estimate (SE) ED visits[b]		Estimate (SE) Acute hospitalizations[c]	
Theme[a]	Track 1	Track 2	Track 1	Track 2	Track 1	Track 2
Theme 1: Physician and clinician lead/medical director engagement with CPC+	0.51	0.58	-13.1**	-3.6	1.4	-1.6
	(0.42)	(0.40)	(5.3)	(5.2)	(2.3)	(2.1)
Theme 2: Co-located teams with care management	0.6	0.71	-1.8	-4.1	-6.2**	2.8
	(0.38)	(0.33)	(5.4)	(5.6)	(2.5)	(2.4)
Theme 3: Offloading routine tasks	0.64	0.74	-6.7	12.4*	-3.7	-5.8**
	(0.30)	(0.27)	(7.8)	(7.2)	(3.0)	(3.0)
Theme 4: Availability and responsiveness	0.64	0.7	-18.6***	1.2	-3.6	-5.3*
	(0.34)	(0.32)	(6.1)	(6.5)	(2.5)	(2.8)
Theme 5: Integrated performance reporting	0.69	0.84	-15.0***	17.4***	-0.4	-0.9
	(0.46)	(0.37)	(5.2)	(5.7)	(2.3)	(3.1)
Theme 6: Identification of patients needing intervention	0.36	0.52	-9.3*	-6.4	-1.0	-2.4
	(0.36)	(0.37)	(5.3)	(5.3)	(2.7)	(2.4)

Source: Mathematica's analysis of Medicare claims data from January 2017 through December 2019 and of the 2017, 2018, and 2019 CPC+ Practice Survey. The CPC+ baseline period corresponds to 2016, Program Year (PY) 1 is 2017, PY 2 is 2018, and PY 3 is 2019. The CPC+ Practice Survey data were not available in 2016. Additionally, data we used to measure Theme 5 (Integrated performance reporting) were not available in 2019.

Note: Each row represents a separate cross-sectional regression model that controls for geographic, average beneficiary, and practice characteristics; prior transformation is not controlled for, however. Beneficiary demographic characteristic controls included average beneficiary age and the proportions for each of the following categories: age, race, sex, original entitlement reason, and dual eligibility. HCC scores and chronic condition controls included average HCC score and proportions for each of the following categories: Tier 4, Tier 5, diabetes, cancer, chronic obstructive pulmonary disease, chronic kidney disease, Alzheimer's and related dementia, heart failure, and behavioral health. Geographic characteristic controls included the median 2014 county-level household income, the 2015 HRR price index, an indicator for a primary care health professionals shortage area, the quartiles of total hospital beds in 2013 per 10,000 2014 county-level total population, the proportions of rural and suburban statuses, the percentage of county-level poverty in 2014, the percentage of county-level Medicare Advantage enrollees/eligible beneficiaries in 2015, and the percentage of people age 25+ with four years of college (2010–2014). The practice characteristic controls included an indicator for ownership by a health system or hospital, categorical counts of primary care practitioners, categorical counts of providers, an indicator for participation in the Shared Savings Program as of January 1, 2017, an indicator for whether the practice has providers from multiple specialties, and the percentage of charges that are for primary care.

[a] Theme scores (and their means) ranged from 0 to 1, with 1 being the highest score possible and indicating the practices used the process or offered the service measured by the theme's CPC+ Practice Survey proxy variable(s) to the fullest extent possible.

[b] Outpatient ED visit utilization, including observation stays, per 1,000 beneficiaries per year during PY 1 to 3.

[c] Acute hospitalizations (short-stay acute care and critical access hospitals), per 1,000 beneficiaries per year during PY 1 to 3.

*/**/*** Significantly different from zero at the 0.10/0.05/0.01 level, two-tailed test.

ED = emergency department; HCC = hierarchical condition category; HRR = hospital referral region; PY = Program Year; SD = standard deviation; SE = standard error.

C. Effects of changes in high-performing practice themes on changes in ED and inpatient utilization

When examining changes in themes and changes in outcomes, there were fewer large significant associations between increases in theme scores and decreases in outpatient ED and inpatient utilization rates than we saw for the cross-sectional results. Changes in availability and responsiveness to patient needs (Theme 4) and identification of patients needing intervention (Theme 6) showed the strongest effect on improving utilization. For Track 1 CPC+ practices, an increase from the minimum to the maximum score (i.e., from 0 to 1) in the changes in availability and responsiveness to patient needs (Theme 4) score was associated with a statistically significant 16.9 visit decrease in the rate of outpatient ED visits per 1,000 beneficiaries (p = 0.01) and an 8.3 visit reduction in the acute hospitalization rate (p = 0.02, Table

5.M.5). Similarly, for Track 1 practices, changes in identification of patients needing intervention (Theme 6) were associated with a lower outpatient ED visit rate by 12.1 visits (p = 0.04) and a lower acute hospitalization rate by 6.9 visits per 1,000 beneficiaries (p = 0.04, Table 5.M.5). These findings are generally consistent with cross-sectional Track 1 estimates' magnitude and direction for those themes (Table 5.M.4).[149]

Furthermore, availability and responsiveness to patient needs (Theme 4) and identification of patients needing intervention (Theme 6) align with the CPC+ care delivery domains of Access and Continuity and Care Management, respectively.

Table 5.M.5. Association between changes in themes from Program Year 1 to 2 (2017–2018) and changes in outcomes (ED visit or acute hospitalization rate) from baseline to Program Year 3 (2016–2019) among CPC+ practices by track

	Mean (SD) Change in theme score from PY 1 to 2		Estimate (SE) Change in ED visits[b]		Estimate (SE) Change in acute hospitalizations[c]	
Theme[a]	Track 1	Track 2	Track 1	Track 2	Track 1	Track 2
Theme 1: Physician and clinician lead/medical director engagement with CPC+	-0.08 (0.44)	-0.05 (0.48)	-5.9 (5.4)	-5.1 (4.9)	1.4 (3.1)	-1.8 (2.7)
Theme 2: Co-located teams with care management	0.15 (0.31)	0.11 (0.29)	-13.5* (8.1)	7.8 (8.9)	-7.3 (4.6)	4.2 (4.8)
Theme 3: Offloading routine tasks	0.12 (0.31)	0.12 (0.31)	9.6 (8.0)	-6.1 (8.0)	-0.7 (4.6)	6.3 (4.3)
Theme 4: Availability and responsiveness	0.06 (0.41)	0.04 (0.39)	-16.9*** (6.5)	-1.4 (6.2)	-8.3** (3.7)	-1.7 (3.3)
Theme 5: Integrated performance reporting	0.14 (0.54)	0.08 (0.43)	1.5 (3.8)	-6.7 (4.4)	1.0 (2.2)	-3.6 (2.4)
Theme 6: Identification of patients needing intervention	0.15 (0.39)	0.14 (0.38)	-12.1** (5.9)	-7.9 (5.7)	-6.9** (3.4)	4.8 (3.1)

Source: Mathematica's analysis of Medicare claims data from January 2016 through December 2019 and of the 2017 and 2018 CPC+ Practice Survey. The CPC+ baseline period corresponds to 2016, Program Year (PY) 1 is 2017, PY 2 is 2018, and PY 3 is 2019. The CPC+ Practice Survey data were not available in 2016.

Note: Each row represents a separate regression model. The changes-in-changes models included practice-level controls. Therefore, we generated the changes-in-changes estimates from models that do not control for the other themes, although they control for practice, geographic, and beneficiary characteristics at the practice level. Beneficiary demographic characteristic controls included average beneficiary age and the proportions for each of the following categories: age, race, sex, original entitlement reason, and dual eligibility. HCC scores and chronic condition controls included average HCC score and proportions for each of the following categories: Tier 4, Tier 5, diabetes, cancer, chronic obstructive pulmonary disease, chronic kidney disease, Alzheimer's and related dementia, heart failure, and behavioral health. Geographic characteristic controls included the median 2014 county-level household income, the 2015 HRR price index, an indicator for a primary care health professionals shortage area, the quartiles of total hospital beds in 2013 per 10,000 2014 county-level total population, the proportions of rural and suburban statuses, the percentage of county-level poverty in 2014, the percentage of county-level Medicare Advantage enrollees/eligible beneficiaries in 2015, and the percentage of people age 25+ with four years of college (2010–2014). The practice characteristic controls included an indicator for ownership by a health system or hospital, categorical counts of primary care practitioners, categorical counts of providers, an indicator for participation in the Shared Savings Program as of January 1, 2017, an indicator for experience with a prior transformation (e.g., participated in MAPCP, medical home recognition, or participated in CPC classic), an indicator for whether the practice has providers from multiple specialties, and the percentage of charges that are for primary care.

[a] Theme scores (and their means) ranged from 0 to 1, with 1 being the highest score possible and indicating the practices used the process or offered the service measured by the theme's CPC+ Practice Survey proxy variable(s) to the fullest extent possible.

[149] Changes-in-changes results were generally more consistent across sensitivity models with and without controls than the cross-sectional results were for sensitivity models with and without controls, which further suggests there is less inherent bias in the changes-in-changes models.

Table 5.M.5. ***(continued)***

[b] Change from baseline (2016) to PY 3 (2019) in outpatient ED visit utilization, including observation stays, per 1,000 beneficiaries per year.

[c] Change from baseline (2016) to PY 3 (2019) in acute hospitalizations (short-stay acute care and critical access hospitals), per 1,000 beneficiaries per year.

*/**/*** Significantly different from zero at the 0.10/0.05/0.01 level, two-tailed test.

ED = emergency department; HCC = hierarchical condition category; HRR = hospital referral region; MAPCP = Multi-payer Advanced Primary Care Practice; PY = Program Year; SD = standard deviation; SE = standard error.

5.M.4 Discussion

A benefit of the Michigan qualitative study design was that it enabled researchers to directly ask high-performing practices to share insights on the techniques, structures, and processes of care they perceived to have contributed to their lower ED and inpatient visit utilization. However, there are limitations with their approach. First, it could be difficult for practices to know the true impacts of those mechanisms since they don't have access to data linking those mechanisms to utilization outcomes. Second, the qualitative study could not control for other characteristics of the practices that could be causally associated with outcomes. Third, their results were based on a small, select sample of practices and might not generalize to all primary care practices. Our analysis of data from the full set of CPC+ practices allows us to overcome some of these limitations of the Michigan study.

Our quantitative results suggest that improvements in practices' availability and responsiveness to patient needs and identification of patients needing intervention have the strongest effects on reducing utilization outcomes, particularly for Track 1 outpatient ED visit utilization. Our results should also be interpreted with caution due to several limitations with our analysis:

1. In addition to the limitations of any self-reported survey data in reliably measuring actual change, given the limited scope of the CPC+ Practice Survey questions, there are aspects of these themes that the CPC+ Practice Survey questions could not capture. For example, the CPC+ Practice Survey proxy variables for Theme 4 (availability and responsiveness to patient needs and patient awareness of the availability) do not fully capture patient awareness of the availability described in the Michigan qualitative study theme, and the survey questions do not capture how quickly practices respond to patient needs.
2. We controlled for many practice-level characteristics in our regression models. However, there could still be factors we cannot measure that both contribute to practices' decreases in outpatient ED visit and acute hospitalization utilization and are related to their decisions to improve on these themes. Such factors would lead us to overestimate the impact of improvements in themes on decreases in these outcomes.[150] For example, the CPC+ Study of Exemplar Practices identified some characteristics that facilitated exemplar practices' ability to implement Medicare acute hospitalization rate reduction strategies, such as organizational support for and staff interest in innovation (Laird et al. 2022, Appendix 4.C). We could not control for this support in our analysis, but it could have influenced practices' performance on both themes and outcomes.
3. Lastly, in the changes-in-changes models, we can only capture the effect of themes that create changes in outcomes in the short run and not the long run.

[150] Alternatively, there could be factors we cannot measure that both contribute to practices' *increases* in outpatient ED visit and acute hospitalization utilization and are related to their decisions to improve on these themes. Such factors would lead us to *underestimate* the impact of improvements in themes on decreases in these outcomes. However, we think such factors would be less common and less of a concern for this analysis than those leading to overestimation (i.e., we think practices could be using many strategies at once to improve utilization).

5.M.5. Conclusion

Changes in availability and responsiveness to patient needs and identification of patients needing intervention were the themes that showed the strongest effects on improving utilization outcomes, particularly for Track 1 ED visit utilization. Findings for these themes were also robust, aligning across our quantitative study's different methods and outcomes. The consistency of the results for these two themes using both the Michigan study's qualitative evidence and our quantitative analyses' evidence also suggests that it is important to invest resources in these care delivery changes. However, further investigation is needed into the remaining themes whose findings were not strongly aligned across the Michigan qualitative study and our quantitative analysis. To better understand whether these themes truly affect ED visit and acute hospitalization outcomes, researchers and model evaluators need better measures of these themes and accurate measurements of them over time. In addition to better measures of these themes, if policymakers want to be able to understand which care processes lead to improved outcomes to inform the return on investment into these aspects of primary care, future innovation models should explicitly test different strategies for changes in care delivery.

5.N. Exemplar quantitative

Key takeaways

The CPC+ exemplar study completed in November 2021 *qualitatively* identified eight activities and four facilitators among CPC+ practices that had the highest probability of achieving substantial reductions in the Medicare acute hospitalization rate (AHR) between 2016 and 2018. We built on this prior work and used *quantitative* methods to test whether there were associations between measurable activities and facilitators and improvements in outcomes among all CPC+ practices.

We found that:

- Improvements in some, but not all, of the exemplar activities and facilitators were associated with decreased AHR.
- Improvements in a few activities also had impacts on outpatient emergency department (ED) visit utilization.
- Improvements in follow-up after hospitalization and ED visits, as well as expanding longitudinal care management, showed the strongest, most robust effects on decreasing the AHR and ED visits.

We did not find evidence that the presence of the facilitators (such as experience and investment in practice transformation and organizational support for and staff interest in innovation) was sufficient for improving outcomes. Due to limitations in our study, our results are best interpreted as revealing aspects of care delivery that are the most promising for further exploration using more rigorous methods.

5.N.1. Introduction

The implementation team conducted a qualitative within- and cross-case comparative analysis of 14 Comprehensive Primary Care Plus (CPC+) practice sites that showed the highest probability of achieving substantial reductions in the Medicare acute hospitalization rate (AHR) between 2016 and 2018 ("exemplars"). The AHR was chosen because it is correlated with costs and also linked to key quality-of-care indicators, including readmission rates and preventable hospitalizations. The team members interviewed the practices from February to December 2020 and asked them what they perceived to be important factors at or outside their practice that explained reductions in the AHR at their practice. The implementation team identified three strategies consisting of two to three activities for reducing AHR, as well as four facilitators that enabled practices to implement the activities (Laird et al. 2022, Appendix 4.C) (Table 5.N.1).

Table 5.N.1. Exemplar strategies, activities, and facilitators

Strategies	Activities
Improve access to primary care	Same-day visits
	Direct access by telephone
	Urgent care sites (system-run)
Expand care management	Follow-up after hospitalization/ED visits
	Longitudinal care management
	Specialized programs
Increase comprehensiveness of care	Breadth of services at practice
	Breadth and depth of care provided by PCP
Facilitators	
Experience and investment in practice transformation	
Using data from CPC+, other payers, health systems, and electronic health record enhancements	
Implementing or enhancing primary care teams through team-based care models	
Organizational support for and staff interest in innovation	

ED = emergency department; PCP = primary care practitioner; CPC+ = Comprehensive Primary Care Plus.

Although the exemplar qualitative study presented valuable evidence of the potential of these activities and facilitators for improving primary care, the methodological approach had some limitations that warranted further investigation. First, the practices likely did not have empirical evidence of the causal mechanisms and pathways leading to lower utilization, so they might have attributed their beneficiaries' utilization to certain activities when other factors could have been responsible. In addition, even if the cited activities helped practices to lower utilization, the activities might only be effective in their particular settings. Therefore, it was unclear whether variation in the implementation of exemplar strategies and the presence of facilitators led to variation in improvements in AHR across all CPC+ practices. The motivation for our quantitative analyses was to determine whether these promising mechanisms of reducing acute care utilization generalize to all CPC+ practices. We quantitively tested these hypothesized mechanisms of action using a methodological approach that included all CPC+ practices and controlled for other observable practice characteristics.

5.N.2. Hypotheses

Based on the exemplar findings, we created a comprehensive list of hypotheses to test the association of each activity or facilitator with changes in AHR. We then selected nine hypotheses for which we identified adequate measures from available data sources to credibly test the exemplar findings.

Due to lack of available measures, we were unable to test hypotheses related to the following activities and facilitators: improving access to primary care through system-run urgent care sites, expanding care management through specialized (i.e., disease-specific) programs, increasing comprehensiveness of care through expanding breadth and depth of care provided by PCP, and implementing or enhancing primary care teams through team-based care models. In prior work, O'Malley, Rich et al. have demonstrated meaningful associations between physician- and practice-level measures of primary care comprehensiveness and cost and utilization outcomes for CPC Classic and CPC+ evaluations (O'Malley et al. 2019, O'Malley et al. 2021).

Table 5.N.2 provides the activity or facilitator identified in the exemplar study, the shorthand for the hypothesis, and the hypothesis.

Table 5.N.2. Hypotheses

Activity or facilitator	Shorthand hypothesis	Hypothesis
Improve access to primary care: Same-day visits	Reserved same-day	CPC+ practices that reserved time specifically for same-day visits in 2018 have larger decreases in AHR than other CPC+ practices.
Improve access to primary care: Direct access by telephone	Telephone access	CPC+ practices that increased timely telephone access to the practice have larger decreases in AHR than other CPC+ practices.
Expand care management: Follow-up after hospitalization/ED	Follow-up	CPC+ practices that increased the rate of follow-up after hospitalization/ED visits have larger decreases in AHR than other CPC+ practices.
Expand care management: longitudinal care management	LCM	CPC+ practices that increased the fraction of patients under longitudinal care management have larger decreases in AHR than other CPC+ practices.
Expand care management: longitudinal care management	Care manager background effect on LCM	For CPC+ practices with care managers with clinical backgrounds in 2018, increasing the number of patients under longitudinal care management results in larger decreases in AHR than doing so in CPC+ practices that had care managers without a clinical background.
Increase comprehensiveness of care: Breadth of services at practice	Additional roles	CPC+ practices that increased their number of service provider roles (e.g., psychologist, referral coordinator, nutritionist, pharmacist) have larger decreases in AHR than other CPC+ practices.
Experience and investment in practice transformation	Larger payments	CPC+ practices with larger new enhanced payments in 2017 have larger decreases in AHR than other CPC+ practices.
Use of data from CPC+, other health systems, and EHR enhancements	Registry data	CPC+ practices that increased the use of registry data to assess or manage groups of patients have larger decreases in AHR than other CPC+ practices.
Organizational support for and staff interest in innovation	Involvement of non-physician staff in CPC+	CPC+ practices in which non-physician staff were more involved in implementing CPC+ in 2017 have larger decreases in AHR than other CPC+ practices.

ED = emergency department; EHR = electronic health record; CPC+ = Comprehensive Primary Care Plus; AHR = Medicare acute hospitalization rate; LCM = longitudinal care management.

5.N.3. Data, sample, and methods

We identified measures for the nine hypotheses using CPC+ Practice Survey data, practice-reported care delivery data submitted to CMS (which we refer to as CDR [care delivery requirement] data), data on CPC+ payments provided by CMS, practice-reported financial data submitted to CMS, and the CPC+ Payer Survey. For each measure (except the larger payments measure), we created scores ranging from 0 to 1.[151] Table 5.N.3 provides the proxy source variables for each measure used to test each of the hypotheses.

Table 5.N.3. Proxy source variables for hypothesis measures

Hypothesis	Proxy source variables
Reserved same-day	Reserved same-day recoded top-box variable: Same-day appointments for patients who need them… 0 = are available only when there are openings for that day, or are generally available by squeezing patients in between scheduled appointments.1 = are generally available through slots reserved for same-day appointments with any physician at this practice site, or are generally available through slots reserved for same-day appointments with the physician who treats them regularly. [PS]
Telephone access	Telephone advice on clinical issues on weekends or after-hours, rescaled pseudo-continuous variable: When patients need it, my practice is able to provide telephone advice on clinical issues on weekends and/or after regular office hours… 0 = never; 0.25 = rarely; 0.5 = sometimes; 0.75 = often; 1 = always. [CDR] Telephone advice on clinical issues during office hours, rescaled pseudo-continuous variable: When patients need it, my practice is able to provide telephone advice on clinical issues during office hours… 0 = never; 0.25 = rarely; 0.5 = sometimes; 0.75 = often; 1 = always. [CDR]
Follow-up	ED outreach, rescaled pseudo-continuous variable: Outreach by this practice site to patients within one week of an ED visit occurs… 0 = for none of this practice's patients; 0.33 = for some of this practice's patients; 0.66 = for many of this practice's patients; 1 = for most or all of this practice's patients. [PS] Hospital outreach, rescaled pseudo-continuous variable: Outreach by this practice site to patients within 3 days of hospital discharge occurs… 0 = for none of this practice's patients; 0.33 = for some of this practice's patients; 0.66 = for many of this practice's patients; 1 = for most or all of this practice's patients. [PS]
LCM	Fraction of patients who are under longitudinal care management. [CDR]
Care manager clinical background	Care manager clinical background. Clinical background of the care managers or care coordinators at this practice site: 1 = Registered nurse (RN); else 0. [PS]
Additional roles	Number of staff roles, rescaled pseudo-continuous variable: Does this practice site have individuals working full-time or part-time in any of the following job roles? Please include all staff who work at this practice site, regardless of who employs them. Clinical psychologist, psychiatrist, or clinical social worker (behavioral health specialists); Referral coordinator or referral specialist (someone who obtains prior authorizations, helps patients obtain appointments with specialists, and/or tracks referrals to specialists); Health educator, dietitian, nutritionist; Clinical pharmacist or doctor of pharmacy. 0 = none; 0.25 = 1 of the roles; 0.5 = 2 of the roles; 0.75 = 3 of the roles; 1 = 4 of the roles. [PS]
Higher payments	New payments. Standardized and Winsorized total new enhanced payment equal to Medicare Care Management Fees (CMFs) plus 2017 CMFs from other payers who said, compared to 2016, they provided additional enhanced payments to support transformation, plus imputed Performance-based Incentive Payments (PBIPs) (PBIP practice received at beginning of 2017 x the average portion retained within their track) per primary care practitioner in the practice. [Data on CPC+ payments provided by CMS, practice-reported financial data submitted to CMS, CPC+ Payer Survey data, and practice-reported roster data submitted to CMS.]
Registry data	Use of registry data to assess or manage groups of patients, rescaled pseudo-continuous variable: A registry is a data system that identifies and tracks patients with specific health conditions, risk states, or medications. At this practice site, registry data to assess or manage care for groups of patients…0 = are not available, 0.33 = are available for 1–2 diseases and/or risk states, 0.67 = are available for 3–5 diseases and/or risk states, 1 = are available for 6 or more diseases and/or risk states. [PS]

[151] For each variable included in the measure, we recoded the variable to be a binary (0 or 1) top-box or rescale to be "pseudo-continuous" (between 0 and 1). For measures that included more than one proxy variable, we then calculated composite scores by taking the mean across all survey items included in the activity for each practice. For the larger payments measure, we put it in standard deviation units.

Table 5.N.3. ***(continued)***

Hypothesis	Proxy source variables
CPC+ involvement of non-physician staff	Clinical support staff involvement in implementing CPC+, rescaled pseudo-continuous variable: Thinking of the different types of staff at this practice site, how involved is each staff type in implementing CPC+? Clinical support staff. 0 = Not at all involved; 0.33 = not very involved; 0.67 = somewhat involved; 1 = very involved. [PS] Administrative support staff involvement in implementing CPC+, rescaled pseudo-continuous variable: Thinking of the different types of staff at this practice site, how involved is each staff type in implementing CPC+? Administrative support staff. 0 = Not at all involved; 0.33 = not very involved; 0.67 = somewhat involved; 1 = very involved. [PS]

CDR = data practices report to CMS on their care delivery requirement activities; CMF = care management fee; CPC+ = Comprehensive Primary Care Plus; ED = emergency department; LCM = longitudinal care management; PBIP = Performance-based Incentive Payment; PS = practice survey.

Table 5.N.4 provides the mean values for each hypothesis measure in Program Year (PY) 1 (2017) and PY 2 (2018) and the change in mean values from PY 1 to PY 2. For some measures, we only used their PY 1 or PY 2 value, either due to availability issues (reserved same-day) or because it made sense to only use the earliest value (higher payments and CPC+ involvement of non-physician staff) or their latest value (care manager clinical background). Some of the measures are fairly topped out (that is, they have high mean values that result in little room for improvement and a lack of variation in the measure), especially the reserved same-day and telephone access measures, which may make it difficult to test the related hypotheses. Overall, practices in both tracks improved on these measures over time.

Table 5.N.4. Trends over time in activity and facilitator measures among CPC+ 2017 Starter Track 1 and 2 practices

	Track 1				Track 2			
Measure	# of Obs.	PY 1	PY 2	Change from PY 1 to PY 2	# of Obs.	PY 1	PY 2	Change from PY 1 to PY 2
Reserved same-day	1,292	n.a.	0.94	n.a.	1,461	n.a.	0.96	n.a.
Telephone access	1,274	0.93	0.97	0.03	1,445	0.94	0.96	0.02
Follow-up	1,287	0.68	0.85	0.17	1,436	0.79	0.88	0.09
LCM	1,253	0.05	0.07	0.02	1,431	0.07	0.07	0.01
Care manager clinical background	1,295	n.a.	0.73	n.a.	1,453	n.a.	0.80	n.a.
Additional roles	1,278	0.31	0.36	0.05	1,413	0.41	0.51	0.10
Higher payments	1,325	-0.23	n.a.	n.a.	1,493	0.56	n.a.	n.a.
Registry data	1,300	0.57	0.62	0.05	1,449	0.71	0.74	0.03
CPC+ involvement of non-physician staff	1,352	0.29	n.a.	n.a.	1,498	0.24	n.a.	n.a.

Source: Mathematica's analysis of PY 1 and PY 2 CPC+ Practice Survey data, P1 and PY 2 practice-reported care delivery data submitted to CMS, PY 1 data on CPC+ payments provided by CMS, PY 1 practice-reported financial data submitted to CMS, PY 1 CPC+ Payer Survey data, and March 2017 practice-reported roster data submitted to CMS.

Note: Measure scores generally ranged from 0 to 1 (except for higher payments), with 1 being the highest score possible and indicating the practices used the process or offered the service to the fullest extent possible as measured by the proxy variable(s) included in the measure. The higher payments measure is standardized to have a mean of 0 and a standard deviation of 1.

CPC+ = Comprehensive Primary Care Plus; PY = Program Year; LCM = longitudinal care management; n.a. = not applicable.

As an initial descriptive check of whether the 14 exemplar practices from the initial qualitative study were indeed exemplars in terms of these hypothesized activities and facilitators, we assessed whether they had higher levels of these measures in PY 1 and larger changes in measures from PY 1 to PY 2 compared to other CPC+ practices nationwide.

To test whether these exemplar activities and facilitators led to greater *improvements* in outcomes among all CPC+ practices, we used two different modeling strategies, which depended on the availability of measures over time. Our preferred approach was to analyze whether changes in these measures predict changes in outcomes (a "changes-in-changes" approach). Specifically, these models tested whether practices with greater improvements in these measures from PY 1 (2017, the earliest available CPC+ Practice Survey and CDR data) to PY 2 also had larger reductions in their adjusted AHR from baseline (2016) to PY 2 (2018). For two hypotheses—larger payments and involvement of non-physician staff in CPC+—we did not calculate the change in the measures from PY 1 to PY 2 because the PY 1 measure scores already implicitly measured changes (there were no CPC+ payments prior to PY 1, and non-physician staff could not be involved in CPC+ before the intervention started). For one hypothesis (reserved same-day), we only had data available for PY 2, so in this case, we used a "quasi-cross-sectional" approach, regressing the changes in AHR on the reserved same-day measure score in PY 2.

All models controlled for the baseline (2016) utilization rate and beneficiary, practice, and geographic factors, while the changes-in-changes models additionally controlled for the PY 1 measure score.

One limitation with the changes-in-changes approach is that there is not much time for the changes in exemplar measures to manifest into changes in outcomes (this is because we can only measure changes in exemplar care process measures across one year starting after the start of CPC+, and our change in outcome measure is only through PY 2). Therefore, in our changes-in-changes regressions, we also assess whether the PY 1 measures predict changes in the adjusted AHR from PY 1 to PY 2. The effects of the PY 1 measure additionally estimate the delayed impacts of prior improvements to these care process measures. However, we interpret these estimates with more caution because these differences might be more likely to reflect other long-standing differences between practices that correlate with trends in outcomes, which could bias our results.

For all the regression analyses, we used the same outcome that was used to identify exemplar practices in the qualitative study: the adjusted change in the practice's average AHR per 1,000 beneficiaries from baseline to PY 2.[152] The AHR was adjusted using a Bayesian multilevel linear regression model that applied both a risk adjustment and a reliability adjustment. In addition, as a secondary outcome, we also used the adjusted change in outpatient emergency department (ED) visits per 1,000 beneficiaries from baseline to PY 2. Because of the short time frame of the analysis, we expected that some of the activities and facilitators might have more immediate effects on outpatient ED visits through diverting care from the ED to primary care.

For both the quasi-cross-sectional model and the changes-in-changes model, we included only 2017 Starter CPC+ practices, analyzed separately by track. We estimated separate regressions for each measure and utilization outcome. These models all used weights accounting for the practice's baseline beneficiary count and clustered standard errors at the practice level.

[152] More specifically, this claims-based measure was the number of hospitalizations at short-stay acute hospitals and critical access hospitals per 1,000 Medicare beneficiaries per year. This AHR measure included emergency department (ED) visits and observation stays if they resulted in an inpatient admission and excluded hospitalizations for elective surgery and planned procedures.

5.N.4. Results

A. Exemplar practices' measure scores

The 14 exemplars had higher average scores for 7 of the 9 hypotheses compared to all other CPC+ practices (Table 5.N.5). These findings validate one aspect of the exemplar qualitative study—that these practices were exceptional in terms of most of these activities and facilitators during PY 1 (or PY 2 if the measure wasn't available in PY 1).

The two measures that were exceptions were longitudinal care management (LCM) and higher payments. For LCM, while the exemplars had lower levels in PY 1, they had much larger increases from PY 1 to PY 2 than all other CPC+ practices (a mean increase of 5 percentage points among exemplars compared to the mean increase of 1 percentage point among all other CPC+ practices). One reason for this might be initial measurement error in early responses to the question. For example, non-exemplars (more than exemplars) might have been confused by the CDR item, due to their initial lack of familiarity with LCM and what it entails, and may have inflated the number of patients they had under LCM in PY 1.

Our findings about payments suggest that, although exemplar practices found payments to be important facilitators, they had lower payments per primary care practitioner (PCP) than other CPC+ practices. This could suggest that in the context of CPC+, the size of the payments was less important than what practices used them for.

Since PY 1 levels of the measures were already very high for exemplars, we found that changes in the measures were generally larger for the other CPC+ practices. The exception to this was LCM, which had the largest potential for growth among both exemplars and other CPC+ practices. The fact that exemplars were close to the maximum value on most measures in PY 1 might indicate that the practices that were most likely to be able to reduce AHR soon after CPC+ started were practices that were already ahead in terms of advanced care processes.

Table 5.N.5. Mean exemplar scores for activity and facilitator measures for exemplar study practices, compared to all other CPC+ practices

	Level in PY 1		Level in PY 2		Change from PY 1 to PY 2	
Measure	CPC+ practices in exemplar study (n = 14)	All other CPC+ practices	CPC+ practices in exemplar study (n = 14)	All other CPC+ practices	CPC+ practices in exemplar study (n = 14)	All other CPC+ practices
Reserved same-day	n.a.	n.a.	1.00	0.95	n.a.	n.a.
Telephone access	1.00	0.94	1.00	0.96	0.00	0.02
Follow-up	0.78	0.74	0.89	0.87	0.11	0.13
LCM	4.45	6.00	9.94	7.11	5.49	1.11
Care manager clinical background	n.a.	n.a.	0.91	0.76	n.a.	n.a
Additional roles	0.56	0.36	0.63	0.44	0.06	0.08
Higher payments	0.01	0.21	n.a	n.a	n.a.	n.a.
Registry data	0.76	0.65	0.66	0.69	-0.10	0.04
CPC+ involvement of non-physician staff	0.31	0.26	n.a.	n.a.	n.a.	n.a.

Table 5.N.5. ***(continued)***

Source: Mathematica's analysis of PY 1 and PY 2 CPC+ Practice Survey data, P1 and the PY 2 practice-reported care delivery data submitted to CMS, PY 1 data on CPC+ payments provided by CMS, PY 1 practice-reported financial data submitted to CMS, PY 1 CPC+ Payer Survey data, and March 2017 practice-reported roster data submitted to CMS.

Note: Sample includes CPC+ 2017 Track 1 and 2 practices combined. Measure scores generally ranged from 0 to 1 (except for higher payments), with 1 being the highest score possible and indicating the practices used the process or offered the service to the fullest extent possible as measured by the proxy variable(s) included in the measure. The higher payments measure is standardized to have a mean of 0 and a standard deviation of 1.

LCM = longitudinal care management; n.a. = not applicable; PY = Program Year.

B. Effects of practice measures on changes in outcomes

We found more favorable and significant effects for the activity measures—particularly those related to care management—than for the facilitator measures. We also found that significant favorable effects were concentrated among the adjusted AHR outcomes, rather than the adjusted outpatient ED visit outcomes.

Reserved same-day. For Track 2 CPC+ practices, we estimated that reserving same-day visits in PY 2 was associated with a statistically significant decrease of 2.8 hospitalizations per 1,000 beneficiaries from baseline to PY 2 compared to practices that did not reserve same-day visits ($p < 0.10$, Table 5.N.6). We found no significant results for Track 1 or on the adjusted change in outpatient ED visit rate. This inconsistency across tracks and outcomes might be because the measure was relatively topped out at 0.94 for Track 1 and 0.96 for Track 2, so there was limited variation we could use to test for impacts for this hypothesis.

Telephone access. For Track 1 CPC+ practices, we estimated that providing full telephone access to clinical advice for all patients in PY 1 (compared to not offering this telephone access to any patients) was associated with a significant decrease of 12.7 hospitalizations per 1,000 beneficiaries ($p < 0.01$) and a significant decrease of 34.3 visits per 1,000 beneficiaries in the adjusted outpatient ED visit rate ($p < 0.01$) from baseline to PY 2.[153] We did not find any significant effects of the *change* in telephone access on the changes in either outcome among CPC+ practices in any track, nor any effects of the level of or change in telephone access on the changes in outcomes among Track 2 CPC+ practices. Similar to our findings for the reserved same-day measure, the inconsistency of results for this hypothesis may relate to how topped out the measure was in PY 1, with a mean of 0.93 for Track 1 and 0.94 for Track 2 CPC+ practices.

Follow-up. For Track 1 CPC+ practices, we estimated that following up with most or all of their patients after ED visits and hospitalizations (compared to following up with none of their patients) in PY 1 was associated with a significant 4.9 decrease in the adjusted AHR ($p < 0.05$) and a significant 7.9 visit decrease in the adjusted outpatient ED visit rate ($p < 0.10$). Improving from following up with no patients after ED visits and hospitalizations to following up with most or all patients from PY 1 to PY 2 was also associated with a statistically significant 4.9 decrease in the AHR ($p < 0.05$). For Track 2 CPC+ practices, we did not find significant impacts on the change in the adjusted AHR. However, we found that following up with most or all patients in PY 1 and improving from not conducting any follow-up in PY 1 to following up to the fullest extent in PY 2, was statistically significantly associated with a respective 9.1 visit decrease ($p < 0.10$) and 14.0 visit decrease ($p < 0.01$) in the adjusted outpatient ED visit rate.

[153] Note that while we report the estimate in units of going from having a score of 0 to a score of 1 for ease of interpretation, this effect is estimated based on the gradation in this measure between 0 and 1. So practices "get credit" for gradations of improvement.

LCM. We found that a one percentage point increase in the proportion of patients that are under longitudinal care management from PY 1 to PY 2 (e.g., from 5% to 6%) was significantly associated with a 0.06 decrease ($p < 0.05$) in the adjusted AHR for Track 1 CPC+ practices, and a 0.09 decrease ($p < 0.01$) in the adjusted AHR for Track 2 practices. For both Tracks, we found no significant effects of the change in LCM on the change in the outpatient ED visit rate, or of the level of LCM in PY 1 on changes in either outcome.

Additional LCM effect if care manager has clinical background. We found no evidence that LCM has an additional effect on changes in AHR and outpatient ED visit outcomes if the care manager in PY 2 has a clinical background.

Additional roles. For CPC+ Track 2 practices, we found that practices' staffing all four possible additional roles (psychologist, referral coordinator, nutritionist, pharmacist) in PY 1 was significantly associated with a 4.5 decrease ($p < 0.01$) in the adjusted AHR compared to practices that did not staff any of these additional roles. Improvements from not having any staff in these additional roles in PY 1 to having all four additional roles in PY 2 was significantly associated with a 2.4 decrease ($p < 0.10$) in the adjusted AHR among Track 2 practices. We found no significant effects of such staffing on the changes in AHR among Track 1 practices, nor on the change in the adjusted outpatient ED visit rate in either track.

Higher payments. For Track 1 CPC+ practices, we found that practices with one standard deviation higher payments per PCP had a significant *increase* of 1.8 ($p < 0.01$) in the adjusted AHR, but otherwise we did not find significant results. Because there is no clear mechanism for more payments to lead to increases in the AHR, it is likely these results are influenced by bias or are spurious due to multiple hypothesis testing.

Registry data. For Track 1 CPC+ practices, improving from having no registry data available in PY 1 to having data available for six or more diseases and/or risk states in PY 2 was associated with a 1.7 decrease ($p < 0.10$) in the adjusted AHR. For Track 2 CPC+ practices, having registry data available for six or more diseases and/or risk states in PY 1 was associated with a 2.5 *increase* ($p < 0.05$) in the adjusted AHR, and going from not having any registry data available to having registry data available for six or more diseases and/or risk states was associated with a 3.8 *increase* in the adjusted AHR ($p < 0.01$). We found no significant effects of registry data levels or changes on changes in outpatient ED visits. Given that the results across tracks are inconsistent, and there is no clear mechanism for improvements in the use of registry data to be associated with increases in the AHR, it is likely these results are influenced by bias or are spurious.

CPC+ involvement of non-physician staff. We found no significant effects of the levels of or changes in the CPC+ involvement of non-physician staff measure on either outcome for Track 1 CPC+ practices. For Track 2 CPC+ practices, we found that having full CPC+ involvement of non-physician staff in PY 1 was associated with a 2.4 *increase* ($p < 0.10$) in the adjusted AHR and a 6.1 visit *increase* ($p < 0.10$) in the adjusted outpatient ED visit rate compared to practices whose non-physician staff were not at all involved in CPC+ in PY 1. Given there is no clear mechanism for CPC+ involvement of non-physician staff to increase the AHR, it is likely these results are due to bias or are spurious.

Based on the results across all the hypotheses, we found three main themes:

1. Most significant results are concentrated on the change in adjusted AHR outcome rather than the change in adjusted outpatient ED visit rate. This may be because exemplar practices were chosen based on their adjusted AHR, and the activity and facilitator exemplar measures were then identified based on discussions about care processes that those practices thought led to the decrease in their AHR, rather than their outpatient ED visit rates.
2. Favorable results (that is, associations between higher PY 1 levels of or increases in exemplar hypothesis measures and decreases in outcomes) were concentrated among hypotheses related to activities rather than facilitators. This may be because the facilitators themselves did not cause changes in the AHR, but merely enabled practices to conduct the activities. This suggests that these facilitators were not sufficient on their own for making important care process changes that would lead to decreases in AHR.
3. Although there were some favorable results for the strategies of improving access to primary care and increasing comprehensiveness of care,[154] the most consistent and largest effects were for the expanding care management strategy.

Table 5.N.6. Association between measures of practices' activities and facilitators and changes in outcomes

	Change in adjusted AHR per 1,000 beneficiaries from baseline to PY 2				Change in adjusted outpatient ED rate per 1,000 beneficiaries from baseline to PY 2			
Measure[a]	Track 1, Effect of level (SE)	Track 1, Effect of change (SE)	Track 2, Effect of level (SE)	Track 2, Effect of change (SE)	Track 1, Effect of level (SE)	Track 1, Effect of change (SE)	Track 2, Effect of level (SE)	Track 2, Effect of change (SE)
Reserved same-day	0.1 (1.3)	n.a.	-2.8* (1.6)	n.a.	3.3 (3.1)	n.a.	-3.5 (3.8)	n.a.
Telephone access	-12.7*** (3.9)	-1.9 (3.8)	-2.3 (3.7)	-5.7 (3.7)	-34.3*** (9.2)	-2.7 (8.9)	-3.8 (8.7)	-13.4 (8.8)
Follow-up	-4.9** (1.9)	-4.9** (1.8)	-0.3 (2.2)	-0.4 (2.2)	-7.9* (4.5)	-3.1 (4.2)	-9.1* (5.2)	-14.0*** (5.1)
LCM	-0.03 (0.03)	-0.06** (0.03)	-0.03 (0.03)	-0.09*** (0.03)	-0.05 (0.07)	-0.08 (0.06)	0.12 (0.08)	0.04 (0.07)
Additional LCM effect if care manager has clinical background	-7.5 (5.7)	-7.5 (4.9)	3.4 (5.5)	5.6 (5.1)	-3.5 (13.4)	-6.1 (11.6)	12.2 (13.0)	4.9 (12.0)
Additional roles	0.6 (1.7)	0.4 (1.4)	-4.5*** (1.4)	-2.4* (1.3)	0.4 (4.0)	4.4 (3.4)	-5.2 (3.4)	2.3 (3.0)
Higher payments	1.8*** (0.6)	n.a.	0.6 (0.4)	n.a.	2.2 (1.5)	n.a.	0.0 (1.0)	n.a.
Registry data	0.8 (1.1)	-1.7* (1.0)	2.5** (1.2)	3.8*** (1.1)	4.7 (2.5)	2.0 (2.3)	-4.3 (2.9)	0.2 (2.5)
CPC+ involvement of non-physician staff	0.2 (1.4)	n.a.	2.4* (1.4)	n.a.	1.6 (3.2)	n.a.	6.1* (3.3)	n.a.

Source: Mathematica's analysis of Medicare claims data from January 2016 through December 2018, PY 1 and PY 2 CPC+ Practice Survey data, P1 and the PY 2 practice-reported care delivery data submitted to CMS, PY 1 data on CPC+ payments provided by CMS, PY 1 practice-reported financial data submitted to CMS, PY 1 CPC+ Payer Survey data, and March 2017 practice-reported roster data submitted to CMS.

[154] Note, however, that our measure of comprehensiveness activities was limited to measuring the number of additional staff roles.

Table 5.N.6. ***(continued)***

Note: This table includes coefficients estimating the association between PY 1 levels of and changes in activities and facilitator measures from PY 1 and PY 2 and changes in outcomes (practice's adjusted average AHR per 1,000 beneficiaries and the adjusted average outpatient ED visit rate per 1,000 beneficiaries). Each row represents a separate regression model, which included practice-level controls. Therefore, we generated the estimates from models that do not control for the other measures listed in the first column, although they control for practice, geographic, and beneficiary characteristics at the practice level. Beneficiary demographic characteristic controls included average beneficiary age and the proportions for each of the following categories: age, race, sex, original entitlement reason, and dual eligibility. HCC scores and chronic condition controls included average HCC score and proportions for each of the following categories: Tier 4, Tier 5, diabetes, cancer, chronic obstructive pulmonary disease, chronic kidney disease, Alzheimer's and related dementia, heart failure, and behavioral health. Geographic characteristic controls included the median 2014 county-level household income, the 2015 HRR price index, an indicator for a primary care health professionals shortage area, the quartiles of total hospital beds in 2013 per 10,000 2014 county-level total population, the proportions of rural and suburban statuses, the percentage of county-level poverty in 2014, the percentage of county-level Medicare Advantage enrollees/eligible beneficiaries in 2015, and the percentage of people age 25+ with four years of college (2010–2014), and HRR fixed effects. The practice characteristic controls included an indicator for ownership by a health system or hospital, categorical counts of primary care practitioners, categorical counts of providers, an indicator for participation in the Shared Savings Program as of January 1, 2017, an indicator for experience with a prior transformation (e.g., participated in MAPCP, medical home recognition, or participated in CPC Classic), an indicator for whether the practice has providers from multiple specialties, and the percentage of charges that are for primary care.

[a] Measures, except for LCM and higher payments, ranged from 0 to 1, with 1 being the highest score possible and indicating the practices used the process or offered the service measured by the source variable(s) included in the measure to the fullest extent possible. The higher payments measure is standardized to have a mean of 0 and a standard deviation of 1. The LCM measure is in percentage units, ranging from 0 to 100.

*/**/*** Significantly different from zero at the 0.10/0.05/0.01 level, two-tailed test.

AHR = acute hospitalization rate; ED = emergency department; HCC = hierarchical condition category; HRR = hospital referral region; LCM = longitudinal care management; MAPCP = Multi-payer Advanced Primary Care Practice; n.a. = not applicable; PY = Program Year; SE = standard error.

5.N.5. Discussion

These quantitative findings build on qualitative findings from our initial exemplar study by controlling for other characteristics of the practices that could be causally associated with outcomes and analyzing the full set of CPC+ practices. Our quantitative results suggest that improvements in follow-up after ED visits and hospitalizations and the percentage of patients under LCM have the strongest effects on reducing utilization outcomes, particularly on the adjusted AHR. Our results should be interpreted with caution due to several limitations with our analysis:

1. In addition to the limitations of any self-reported survey data in reliably measuring actual change, given the limited scope of the CPC+ Practice Survey and CDR questions, there are aspects of the exemplar activities and facilitators that our data sources could not capture. Indeed, many of the measures were topped out already in PY 1, and likely did not capture important facets of the activities that the exemplars did capture.
2. We controlled for many practice-level characteristics in our regression models. However, there could still be factors we cannot measure that both contribute to practices' decreases in the adjusted AHR and the adjusted outpatient ED visit rate and are related to their decisions to improve on these measures. Such factors would lead us to overestimate the impact of improvements in exemplar measures on decreases in these outcomes. Alternatively, there could be factors we cannot measure that both contribute to practices' *increases* in the AHR and outpatient ED visit rate and are related to their decisions to improve on these measures. Such factors would lead us to *underestimate* the impact of improvements in measures on decreases in these outcomes. Given that we found some unfavorable statistically significant effects, it is possible that this issue impacted at least some of our hypotheses.

3. We only capture the effect of changes made from mid-PY 1 to PY 2 in the measures that created changes in outcomes through PY 2 and not in the long run.[155]
4. Lastly, due to the multiple outcomes, tracks, and practice activity and facilitator measures that we tested, we expect that there could be some statistically significant results just by chance. For this reason, we focus on results that show some consistency across tracks and outcomes.

These limitations could contribute to why we found results for the three facilitators that go against our hypotheses. Given that there are several known limitations, results should be interpreted as revealing aspects of care delivery that are most promising for further exploration, and suggesting that improvements to care management might be the most promising.

5.N.6. Conclusion

Based on our analysis, improvements in the follow-up after ED visits and hospitalizations and LCM measures showed the strongest effects on improving utilization outcomes. We also found evidence that reserving same-day visits, having telephone access, and increasing the breadth of services at the practice by increasing the number of service provider roles may lead to improvements in AHR. We did not find evidence that facilitators led to better outcomes on their own. Given the limitations of our study, further investigation is needed to better understand whether these activities truly affect ED visit and acute hospitalization outcomes. To do so, researchers and model evaluators need better measures of these activities and accurate measurements of them over time, including measurements before the model starts. Survey items should be designed to measure care process activities explicitly and concretely, with ample room for growth for the participating population of practices. In addition to using better measures of activities, future innovation models should explicitly test different strategies for changes in care delivery—for example, by randomly assigning care delivery process activities for participating practices to pursue. Such approaches could help policymakers understand the care processes that lead to improved outcomes, as well as the return on investments into these aspects of primary care.

[155] As a follow-up, Appendix 5.O. Synthesis describes results from regressions of changes in promising care processes from PY 1 to PY 2 on changes in outcomes from PY 2 through PY 5.

5.O. Synthesis

Key takeaways

We tested whether improvements in select primary care processes among CPC+ practices were associated with decreases in outpatient emergency department (ED) visits and acute hospitalizations from baseline to follow-up periods in Program Year (PY) 2 through PY 5. The most robust evidence we found was that practice-reported improvement in the availability for same-day and next-day visits was associated with decreases in both ED visits and acute hospitalizations among CPC+ beneficiaries. Due to limitations in our study, our results are best interpreted as revealing an aspect of care delivery that is promising for further exploration using more rigorous methods.

5.O.1. Introduction

Over the intervention period, CPC+ reduced both outpatient ED visits and acute hospitalizations in both tracks. In addition, the implementation analysis found that CPC+ practices improved across a variety of primary care delivery processes. In the synthesis work, we aimed to connect improvements in care process measures (that is, increases) and improvements in Medicare fee-for-service (FFS) utilization outcomes (that is, decreases) among practices participating in the model to better understand the mechanisms through which CPC+ impacted these two key outcomes.

Based on the CMS implementation guide and related literature, we identified two potential mechanisms through which reductions in these outcomes might occur: (1) improved patient health status by better management of chronic illness and behavioral health conditions ("preventive pathway"), and (2) improved access influencing patients' choices about where to seek care after symptoms develop for conditions that can be treated in a primary care setting ("patient choices pathway").

We have conducted several earlier analyses that helped test these two pathways using data through PY 2 (2018) or PY 3 (2019). These are summarized in other appendices to this report:

- Appendix 5.L contains the results of investigating the role of the patient choices pathway in reducing outpatient ED visits through PY 3, by testing whether (1) access to primary care improved over time among CPC+ practices, (2) CPC+ reduced primary care substitutable (PCS) ED visits on weekdays and non-weekdays, and (3) greater changes in access were associated with fewer PCS ED visits among CPC+ practices. We measured access using items from CPC+ Practice Surveys and practice-reported care delivery data submitted to CMS (which we refer to as CDR [care delivery requirement] data), which were consistently reported from PY 1 through PY 3. We found that practice-reported access to primary care modestly improved among CPC+ practices during the intervention period, and some of these improvements were associated with lower levels of utilization. Specifically, increases in same- or next-day appointment availability from PY 1 to PY 2 were significantly associated with decreased PCS ED visit use for Track 1 only, and increases in care manager (CM) full time equivalents (FTEs) per 1,000 patients were significantly associated with decreased PCS ED visit use for Track 2 only. However, we found that the improvements that we can measure with Practice Survey and CDR data cannot explain the small effect of CPC+ on PCS ED visits.

- Appendix 5.M contains the results of a quantitative analysis among all CPC+ practices to test whether practice identification themes qualitatively identified by the "CPC+ Michigan ED and Inpatient Utilization High-Performing Practice Study" (Finkel and Marriott 2021) led to decreases in outpatient ED visit and acute hospitalization utilization from baseline through PY 3. We measured practice identification themes using related items from the Practice Surveys and found that improvements in practices' availability and responsiveness to patient needs (that is, offering same- or next-day appointments and after-hours access) and identification of patients needing intervention (that is, using a standard, integrated method to stratify patients by risk level and registry data to identify and manage groups of patients) showed the strongest, most robust effects on improving these outcomes.
- Appendix 5.N contains the results from a follow-up on the CPC+ Exemplar study that *qualitatively* identified eight activities and four facilitators among CPC+ practices that had the highest probability of achieving substantial reductions in the Medicare acute hospitalization rate (AHR) between baseline and PY 2 (Laird et al. 2022, Appendix 4.C). We used *quantitative* methods to test whether there were associations between measurable activities (from the CPC+ Practice Surveys and CDR data) or facilitators and improvements in outcomes among all CPC+ practices. We found that improvements in follow-up after hospitalizations and ED visits, as well as expanding longitudinal care management (LCM), showed the strongest, most robust effects on decreasing the AHR and ED visits.[156]

In this Appendix, we explore further the care process measures that showed the most promise from our earlier investigations.[157] For our main analysis, we looked at how changes in these promising care process measures from PY 1 to PY 2 (where practices demonstrated the largest gains) affected changes **from baseline to follow-up periods PY 2 through PY 5** in the two main outcomes of interest: outpatient ED visits and acute hospitalizations.

5.O.2. Data, sample, and methods

To test the associations between promising care processes and outcomes, we used regression models similar to those reported in Appendices 5.L, 5.M, and 5.N. Specifically, we regressed changes in utilization outcomes on changes in care process measures, while controlling for important practice characteristics. We used the sample of all 2017 Starter CPC+ practices that had care measures available in PY 1 (2017) and PY 2 (2018). The rest of this section provides more details on the care process measures, utilization outcomes, model, and limitations to our analysis.

A. Primary care process measures

The primary care process measures we tested are based on those that were found to have significant favorable results on outcomes and were reported in Appendix 5.L. Patient Choices Pathway, Appendix 5.M. Michigan Quantitative, or Appendix 5.N. Exemplar Quantitative. Table 5.O.1 lists the names of the care process measures, the descriptions of and sources for them, which pathways we theorize the care processes work through, and the appendices in which we report favorable results for the care process

[156] The CPC+ Exemplar study focused on activities and facilitators that had the highest probability of improving just the AHR, but for the Exemplar quantitative study, we examined the association of activities and facilitators with *both* the AHR and outpatient ED visits.

[157] The Michigan quantitative and Exemplar quantitative studies were not limited to care processes that went through the preventive and patient choices causal pathways (that is, that were related to access or care management), but the care processes that showed the most promise in these studies were under the access and care management functions.

measure using data through PY 3. All the measures are based on the CPC+ Practice Surveys or CDR data. For the follow-up, identification of patients for intervention, and same-day visits measures, we rescaled each variable included in the measure to be "pseudo-continuous" with values between 0 and 1. We then calculated composite scores by taking the mean across all survey items included in the activity for each practice. The LCM measure is in percentage units (percentage of patients enrolled in LCM across all risk tiers), and the *CMs per patient measure* is in FTEs per 1,000 patient units.

Table 5.O.1. Care process measures

Care process	Description [data source]	Theorized pathway	Original appendix source
Follow-up	Follow-up after ED within one week and after hospitalizations within 3 days, on scale of 0 to 1 [PS]	Preventive	5.N. Exemplar quantitative
LCM	Percentage of patients who are enrolled in LCM, on scale of 0 to 100 [CDR]	Preventive	5.N. Exemplar quantitative
Identification of patients for intervention	Use of patient stratification by risk level for all patients, use of registry data to assess or manage groups of patients, on scale of 0 to 1 [PS]	Preventive	5.M. Michigan quantitative
CMs per patient	Care manager FTEs [PS] per 1,000 patients [CDR]	Preventive and patient choices	5.L. Patient choices pathway
Same-day visits	Availability of same-day or next-day visits with patients who need them, on scale of 0 to 1 [PS and CDR]	Patient choices	5.L. Patient choices pathway

CDR = data practices report to CMS on their care delivery requirement activities; CM = care manager; ED = emergency department; FTE = full time equivalent; LCM = longitudinal care management; PS = Practice Survey.

Table 5.O.2. provides the mean values for each care process measure in PY 1 and PY 2, the percent average change in values from PY 1 to PY 2, and the standard deviation in the change for both Track 1 and Track 2 practices. We found that practices in both tracks improved in these measures from PY 1 to PY 2, but the amount, and the degree of variation, differed by the measure.

Table 5.O.2. Trends over time in average care process measures among CPC+ 2017 Starter Track 1 and 2 practices

	Track 1				Track 2			
Measure	PY 1	PY 2	Percent change from PY 1 to PY 2	SD of change from PY 1 to PY2	PY 1	PY 2	Percent change from PY 1 to PY 2	SD of change from PY 1 to PY 2
Follow-up	0.68	0.85	25	0.27	0.79	0.88	11	0.23
LCM	6.0	8.1	35	17.3	6.3	7.2	13	15.1
Identification of patients for intervention	0.25	0.39	56	0.39	0.41	0.55	37	0.39
CMs per patient	0.27	0.36	33	0.79	0.28	0.33	18	0.73
Same-day visits	0.90	0.93	3	0.14	0.93	0.94	1	0.13

Source: Mathematica's analysis of PY 1 and PY 2 CPC+ Practice Survey data, and P1 and PY 2 practice-reported care delivery data submitted to CMS.

Note: The follow-up, identification of patients for intervention, and same-day visit measures ranged from 0 to 1, with 1 being the highest score possible and indicating the practices used the process or offered the service to the fullest extent possible as measured by the variable(s) included in the measure. The LCM measure is on a scale of 0 to 100 as the percentage of patients enrolled in LCM across all risk tiers. The CMs per patient measure is in units of FTEs per 1,000 patients.

CM = care manager; LCM = longitudinal care management; PY = Program Year; SD = standard deviation.

B. Outcomes

The primary outcomes for this analysis were changes in outpatient ED visits per 1,000 beneficiaries and acute hospitalizations per 1,000 beneficiaries from baseline to follow-up periods in PY 2 through PY 5. We also examined more specific outcomes that may capture visits that are more likely to be impacted by the theorized pathways: PCS ED visits for the patient choices pathway, and medical acute hospitalizations and medical acute hospitalizations without a complication or comorbidity (CC) or a major complication or comorbidity (MCC) for the preventive pathway. Medical acute hospitalizations are on average lower acuity than surgical hospitalizations, and medical acute hospitalizations without a CC or MCC are on average lower acuity than all medical acute hospitalizations. We hypothesize that medical hospitalizations and the categories that are lower acuity will be more likely to be prevented by CPC+ (see Appendix 5.J.).

We constructed PCS ED visits and preventable ED visits as mutually exclusive subsets of the outpatient ED visits and align with the New York University Emergency Department Algorithm (Johnston et al. 2017). We calculated PCS ED visits as the sum of probabilities that the visit was (1) nonemergent or (2) emergent but treatable in a primary care setting. We calculated preventable ED visits as (1) emergent/ED care required but preventable or (2) avoidable if appropriate ambulatory care had been given (see Appendix 5.C for more details). We constructed medical acute hospitalizations and medical acute hospitalizations without MCC/CC or co-morbidities using the Medicare severity diagnosis-related group (see Appendix 5.J. for more details).

C. Modeling strategy

The models tested whether practices with greater improvements in the care process measures from PY 1 (the earliest available CPC+ Practice Survey and CDR data) to PY 2 also had larger reductions in their outcomes from baseline to each follow-up period, PY 2 through PY 5.

All models controlled for the baseline utilization rate and PY 1 care process measure score, as well as beneficiary, practice, and geographic factors, including hospital referral region (HRR) fixed effects.[158] Specifically, these included the following:

Beneficiary demographic characteristic controls included average beneficiary age and the proportions for each of the following categories: age, race, sex, original entitlement reason, and dual eligibility. Risk controls included average hierarchical condition category (HCC) score and proportions for the following categories: Tier 4 and Tier 5 beneficiaries (defined based on payment attribution methodology), diabetes, cancer, chronic obstructive pulmonary disease, chronic kidney disease, Alzheimer's and related dementia, heart failure, and behavioral health.

Geographic characteristic controls included an indicator for a primary care health professionals shortage area, indicators for the practice being rural or suburban, the percentage of county-level poverty in 2014, quartiles for total hospital beds in 2015 per 10,000 population in county, and the percentage of county-level Medicare Advantage enrollees/eligible beneficiaries in 2015.

[158] One difference from the model used in this Appendix and the one used in Appendix 5.L Patient Choices Pathway and 5.M Michigan Quantitative is we that included HRR fixed effects to help control for effects COVID-19 may have had in PY 4 and PY 5. We also included HRR fixed effects even in models for PY 2 and PY 3 in this Appendix so that we would have a consistent set of controls across all years.

Practice characteristic controls included an indicator for ownership by a health system or hospital, categorical counts of primary care practitioners, categorical counts of providers, the number of attributed Medicare beneficiaries, the number of Medicare beneficiaries per primary care practitioner, an indicator for participation in the Shared Savings Program as of January 1, 2017, an indicator for experience with a prior transformation initiative (e.g., participated in the Multi-Payer Advanced Primary Care Practice model, medical home recognition, or participated in CPC Classic), an indicator for whether the practice has providers from multiple specialties, an indicator for whether the practice has any nurse practitioners or physician assistants, an indicator for being an Indian Health Center, and the percentage of charges that are for primary care.[159]

As with our main impact analysis, we included only 2017 Starter CPC+ practices and analyzed results separately by track. Unlike the main impact analysis, we only included CPC+ practices and not comparison practices. We estimated separate regressions for each measure, utilization outcome, and follow-up year. These models all used weights accounting for the practice's baseline beneficiary count and clustered standard errors at the practice level.

D. Limitations

Our results should be interpreted with caution due to several limitations with our analysis:

1. We focused our analyses on care processes that we could measure with available data and we had previously found promising, which might have missed care processes that are important for determining outcomes. Although the initial analyses covered a variety of care processes—those that we theoretically believed should be related to utilization outcomes, as well as those that were identified through qualitative analyses to be related to key outcomes—it is still possible we missed important care processes. These care processes we might have missed include those that we did not have adequate data to measure and those that were not identified as part of our theoretical models or mentioned in qualitative studies but were still important for changing outcomes. In addition, since we only followed up on *promising* care processes, we may have missed some care processes that had no immediate impacts but had longer-term effects.
2. In addition to the limitations of any self-reported survey data in reliably measuring actual change, given the limited scope of the CPC+ Practice Survey and CDR questions, there are some aspects of the care delivery processes that our data sources could not capture. For example, in the qualitative study of exemplar practices (Laird et al. 2022, Appendix 4.C), interviews suggested that the methodology for increasing same-day and next-day visits was important for increasing access (that is, hiring staff to provide same-day visits versus adding same-day visit slots to existing practitioners' schedules). While in the data we were able to measure the rough portion of patients that practices were able to provide same and next visits for, we were not able to measure consistently over time the way these visits were reserved.
3. We did not have measures of care processes prior to the start of CPC+ in 2017. We used the earliest available data from the PY 1 Practice Survey (administered from May 2017 through September 2017) and PY 1 Q1 and Q2 responses from the CDR. However, practices may have already made changes to

[159] We define proportion of charges that are for primary care as the proportion of charges that are for office visit evaluation and management, nursing home and home care, welcome to Medicare and annual wellness visits, advance care planning, chronic care management services, and transitional care management services among National Provider Identifiers with a primary care specialty, according to SK&A.

their care processes prior to our measurements, which could understate the relationship between improvements in care processes and utilization measures in our models.

4. We controlled for many practice-level characteristics in our regression models. However, there could still be factors we cannot measure that contribute to practices' decreases in the AHR and outpatient ED visits and are related to their decisions to improve on these measures. Such factors would lead us to overestimate the impact of improvements in care process measures on decreases in these outcomes. Alternatively, there could be factors we cannot measure that contribute to practices' *increases* in the AHR and outpatient ED visit rate and are related to their decisions to improve on these measures. Such factors would lead us to *underestimate* the impact of improvements in measures on decreases in these outcomes.
5. Lastly, due to the multiple outcomes, tracks, and care process measures tested, we expect that there could be some statistically significant results just by chance. For this reason, we focused on results that showed some consistency across tracks, outcomes, and years.

5.O.3. Results

We found the largest favorable associations (that is, improvements in care process measures being related to decreases in utilization measures) for the availability of same-day and next-day visits measure. We found some evidence for unfavorable effects of the LCM and CMs per patient measures, but these were sensitive to changes in our model specification and thus are likely due to bias.

Follow-up. For Track 1 and 2 CPC+ practices, we found no statistically significant effects of improving follow-up on changes in outpatient ED visits in any year (Table 5.O.3.). We estimated that improving from following up with no patients after ED visits and hospitalizations to following up with most or all patients from PY 1 to PY 2 was associated with a statistically significant decrease in the AHR of 10.5 per 1,000 beneficiaries from baseline to PY 5 for Track 1 practices ($p < 0.10$) and a 17.5 per 1,000 beneficiaries decrease for Track 2 practices ($p < 0.05$).

LCM. We found no significant effects of improvements in LCM for Track 1 practices. For Track 2 practices, we found that a one percentage point increase in the proportion of patients across all risk tiers that are enrolled in LCM from PY 1 to PY 2 (for example, from 5 to 6 percent) was significantly associated with an increase in the outpatient ED visit rate of 0.3 per 1,000 beneficiaries ($p < 0.05$) from baseline to PY 4, and a 0.6 increase ($p < 0.01$) from baseline to PY 5.

Identification of patients for intervention. We found no significant effects of improvements in identification of patients for intervention in either track on either outcome in any year.

CMs per 1,000 patients. For Track 1 practices, we found increasing the number of CM FTEs per 1,000 patients by 1 from PY 1 to PY 2 was associated with increasing the AHR by 4.3 ($p < 0.10$) from baseline to PY 2, and by 5.2 ($p < 0.10$) from baseline to PY 3. For Track 2 practices, we found it was associated with increasing the AHR by 6.4 ($p < 0.05$) from baseline to PY 3, by 8.5 ($p < 0.01$) from baseline to PY 4, and by 5.3 ($p < 0.10$) from baseline to PY 5.

Same-day visits. For Track 1 practices, we found that improving from not having same-day and next-day visits available for any patients to having such visits available for all patients who needed it was associated with a 41.8 decrease ($p < 0.05$) in the outpatient ED visit rate from baseline to PY 2, a 35.0 decrease ($p < 0.10$) in the outpatient ED visit rate from baseline to PY 3, and a 45.9 decrease ($p < 0.01$)

in the outpatient ED visit rate from baseline to PY 4. For Track 1 practices, we found that improving the same-day visit measures was associated with a 25.2 decrease ($p < 0.05$) from baseline to PY 2 in the AHR, and a 27.2 decrease ($p < 0.05$) from baseline to PY 3. For Track 2 practices, we found no significant effect of improving the same-day visit measure on the outpatient ED visit rate, but we found that going from no availability to availability for all patients who need it was associated with a 19.2 decrease ($p < 0.10$) in the AHR from baseline to PY 3 and a 28.5 decrease ($p < 0.01$) in the AHR from baseline to PY 5.

Table 5.O.3. Associations of changes in care process measures and changes in outpatient ED visits and acute hospitalizations over the CPC+ intervention period

	Outpatient ED visits per 1,000 beneficiaries				Acute hospitalizations per 1,000 beneficiaries			
Change from baseline to:	PY 2 Coef (SE)	PY 3 Coef (SE)	PY 4 Coef (SE)	PY 5 Coef (SE)	PY 2 Coef (SE)	PY 3 Coef (SE)	PY 4 Coef (SE)	PY 5 Coef (SE)
Follow-up								
Track 1	-1.3	13.3	2.4	5.1	-3.8	0.3	0.6	-10.5*
	(11.3)	(12.0)	(10.9)	(11.9)	(6.3)	(6.8)	(6.1)	(6.3)
Track 2	-7.6	-6.8	-12.5	-6.5	-3.3	-9.1	-7.2	-17.5**
	(12.6)	(13.3)	(12.0)	(13.0)	(7.2)	(7.4)	(6.8)	(7.0)
LCM (1pp)								
Track 1	-0.1	0.2	0.2	0.2	0.0	0.1	0.0	-0.1
	(0.2)	(0.2)	(0.2)	(0.2)	(0.1)	(0.1)	(0.1)	(0.1)
Track 2	0.2	0.0	0.3**	0.6***	-0.1	-0.1	-0.1	0.0
	(0.2)	(0.2)	(0.2)	(0.2)	(0.1)	(0.1)	(0.1)	(0.1)
Identification of patients for intervention								
Track 1	1.7	-5.9	1.7	-0.1	-1.9	-4.4	-3.3	-3.5
	(5.9)	(6.2)	(5.7)	(6.2)	(3.2)	(3.5)	(3.2)	(3.3)
Track 2	2.0	-0.3	2.6	7.1	0.2	1.8	4.9	4.0
	(5.7)	(6.0)	(5.4)	(5.9)	(3.2)	(3.3)	(3.1)	(3.2)
CMs per 1,000 patients								
Track 1	3.8	4.3	2.8	4.9	4.3*	5.2*	-1.2	2.3
	(4.5)	(4.8)	(4.4)	(4.8)	(2.5)	(2.7)	(2.4)	(2.5)
Track 2	-3.4	-7.9	-6.4	6.3	2.6	6.4**	8.5***	5.3*
	(5.2)	(5.5)	(5.0)	(5.4)	(3.0)	(3.1)	(2.8)	(2.9)
Same-day visits								
Track 1	-41.8**	-35.0*	-45.9***	-27.0	-25.2**	-27.2**	-6.1	-13.6
	(18.1)	(19.5)	(17.6)	(19.2)	(10.0)	(10.9)	(9.8)	(10.1)
Track 2	-24.4	10.5	23.3	-1.9	-18.3	-19.2*	-2.8	-28.5***
	(19.6)	(20.7)	(18.6)	(20.3)	(11.2)	(11.5)	(10.6)	(11.0)

Source: Mathematica's analysis of Medicare claims data from January 2016 through December 2021; 2017 and 2018 CPC+ Practice Survey; and 2017 Q1–Q2 and 2018 Q3–Q4 CPC+ Practice Portal data.

Note: The follow-up, identification of patients for intervention, and same-day visit measures ranged from 0 to 1, with 1 being the highest score possible and indicating the practices used the process or offered the service to the fullest extent possible as measured by the variable(s) included in the measure. The LCM measure is on a scale of 0 to 100 as the percentage of patients enrolled in LCM across all risk tiers. The CMs per patient measure is in units of FTEs per 1,000 patients. Each row represents a separate regression model that included practice-level controls. Therefore, we generated these estimates from models that do not control for the other care process measures, although they control for practice, geographic, and beneficiary characteristics at the practice level. These controls include HRR fixed effects to help control for effects COVID-19 may have had in PY 4 and PY 5. See methods section for list of controls.

*/**/*** Significantly different from zero at the 0.10/0.05/0.01 level, two-tailed test.

CM = care manager; Coef = coefficient; ED = emergency department; FTE = full-time equivalent; HRR = hospital referral region; LCM = longitudinal care management; SE = standard error; pp = percentage point; PY = Program Year.

A. Sensitivity tests

The unfavorable effects we found of improvements in LCM and CMs per 1,000 patients being associated with increases in utilization were not robust to sensitivity checks, but the favorable effects of improvements in same-day visits were more robust.

The unfavorable effects we found for some care process measures, which we do not have a strong causal mechanism for explaining, may suggest that our results are influenced by bias. To explore this, we performed some additional sensitivity tests on the care process measures where we found significant unfavorable effects (LCM and CMs per 1,000 patients) and the measure where we found significant, consistently favorable effects (same-day visits).

The two standard sensitivity tests we conducted were:

1. Including additional controls for changes in the average HCC score, age, and dual status of beneficiaries assigned to the practice between baseline and the relevant follow-up period. This helped control for changes in composition of beneficiaries (which we refer to as "composition controls") at the practice that could drive changes in outcomes.
2. Analyzing the effects of changes in the care process measure from PY 1 to PY 3 on changes in outcomes from baseline to PY 3 and onward. Although most of the average improvement in care process measures happened between PY 1 and PY 2, there were some subsequent improvements. If we were to find consistent results using this alternate measure, that would give more credence to our results.

For CMs per 1,000 patients, we did one additional sensitivity test that looked at the effects of changes in CMs per *100 Medicare beneficiaries* rather than per *1,000 patients.*[160] Since CPC+ was designed to be a practice-wide initiative, we expected CMs per 1,000 patients to be the most relevant measure. However, the number of patients has very large positive and negative percentage changes over time, whereas the number of Medicare beneficiaries in a practice has on average lower magnitude percentage changes.[161] It is possible that, since the number of patients is reported by the practice whereas the number of Medicare beneficiaries is based on attribution using Medicare FFS claims, the large percentage changes in patient population are driven more by noise than by Medicare FFS beneficiaries. We therefore tested whether the unfavorable results persisted using CMs per 100 Medicare beneficiaries as an alternative measure.

LCM. We found that the Track 2 unfavorable results did not change considerably with the addition of changes in composition controls, but they were eliminated when we instead assessed the association of the PY 1 to PY 3 change in LCM on changes in outcomes (Table 5.O.4). This suggests that the unfavorable results we originally found are not robust and meaningful. We found one new, significant unfavorable result using the PY 1 to PY 3 change in LCM, but it was not present in any of the other specifications, suggesting it is not a robust finding either.

[160] To keep the magnitude of the CM measure relatively similar between the per-patient and per-beneficiary measures, for the per-beneficiary measure, we calculated it as per 100 rather than per 1,000 (the median number of patients per practice is around 4,000, whereas the median number of Medicare beneficiaries is around 400).

[161] The 10th percentile for the distribution of percentage changes in patients from PY 1 to PY 2 was -32 percent and the 90th percentile was 46 percent. The 10th percentile for the distribution of percentage changes in Medicare beneficiaries from PY 1 to PY 2 was -14 percent and the 90th percentile was 24 percent.

CMs per patient. For Track 1, the significant unfavorable effects of improving CMs per patient on changes in the **AHR** from baseline to PY 2 and baseline to PY 3 are not present for any of the sensitivity tests, largely due to an attenuation of these effects under these alternate specifications. For Track 2, the different sensitivity tests lead to attenuation and elimination of the significant unfavorable results, and some isolated favorable results. This suggests that the unfavorable results we found of improvements in CMs per patients may have been biased or spurious. We also found one new unfavorable significant association when we used the PY 1 to PY 3 change in CMs per patient; however, this effect is not present in any of the other specifications, suggesting it is not a robust finding either.

Same-day visits. For Track 1, including changes in composition controls attenuated the original significant favorable effects by around 9 to 18 percent, resulting in the effects on changes in the **outpatient ED visit rate** in one year becoming statistically non-significant. Using the PY 1 to PY 3 change in same-day visits eliminated all of the significant Track 1 effects of improvements in the same-day visit measure.[162] For Track 2, including changes in composition controls also led to some attenuation of the significant favorable effects on the **AHR,** and the creation of one significant unfavorable effect. Using the PY 1 to PY 3 change led to finding a significant favorable effect of same-day visits on decreases in the **AHR** from baseline to PY 4 rather than from baseline to PY 3. Although the results for same-day visits were somewhat sensitive to the specification, we still found some favorable results for each specification. Overall, we found more favorable than unfavorable results.

Table 5.O.4. Sensitivity of associations of changes in select care process measures and changes in outpatient ED visits and acute hospitalizations over the CPC+ intervention period

	Outpatient ED visits per 1,000 beneficiaries				Acute hospitalizations per 1,000 beneficiaries			
Change from baseline to:	PY 2	PY 3	PY 4	PY 5	PY 2	PY 3	PY 4	PY 5
LCM								
Track 1								
Main	-0.1	0.2	0.2	0.2	0.0	0.1	0.0	-0.1
Change in composition controls[a]	-0.1	0.2	0.1	0.1	-0.1	0.1	-0.1	-0.1
PY 1 to PY 3 change[b]	n/a	0.0	0.0	-0.2	n/a	0.0	0.0	-0.1
Track 2								
Main	0.2	0.0	0.3**	0.6***	-0.1	-0.1	-0.1	0.0
Change in composition controls[a]	0.2	0.0	0.3**	0.5***	-0.1	-0.1	-0.1	-0.1
PY 1 to PY 3 change[b]	n/a	-0.5*	-0.1	0.1	n/a	0.0	0.3**	0.1

[162] One reason using the PY 1 to PY 3 change may have eliminated the Track 1 results is that we found when we controlled for the PY 1 level of same-day visits, there was less variation in the change from PY 1 to PY 3 than there was in the change from PY 1 to PY 2. Alternatively, for Track 2, when we controlled for the PY 1 level, there was more variation in the change in same-day visits from PY 1 to PY 3 than there was in the change from PY 1 to PY 2.

Table 5.O.4. ***(continued)***

Change from baseline to:	Outpatient ED visits per 1,000 beneficiaries				Acute hospitalizations per 1,000 beneficiaries			
	PY 2	PY 3	PY 4	PY 5	PY 2	PY 3	PY 4	PY 5
CMs per patient								
Track 1								
Main	3.8	4.3	2.8	4.9	4.3*	5.2*	-1.2	2.3
Change in composition controls[a]	2.6	3.4	1.8	2.0	3.2	3.1	-3.4	-1.6
PY 1 to PY 3 change[b]	n/a	-0.6	4.2	7.8*	n/a	-0.7	2.5	2.0
CMs per 100 beneficiaries[c]	-1.7	1.4	1.0	-1.0	2.1	1.5	0.7	0.4
Track 2								
Main	-3.4	-7.9	-6.4	6.3	2.6	6.4**	8.5***	5.3*
Change in composition controls[a]	-3.2	-8.8*	-7.5	3.5	0.6	3.8	5.5**	1.9
PY 1 to PY 3 change[b]	n/a	-2.3	-1.2	2.2	n/a	6.0**	3.3	1.4
CMs per 100 beneficiaries[c]	-5.8	-12.7**	-11.3**	2.1	0.5	1.6	-2.0	-1.8
Same-day visits								
Track 1								
Main	-41.8**	-35.0*	-45.9***	-27.0	-25.2**	-27.2**	-6.1	-13.6
Change in composition controls[a]	-35.4**	-28.7	-39.0**	-21.6	-22.6**	-24.7**	-2.9	-11.1
PY 1 to PY 3 change[b]	n/a	4.1	1.4	6.3	n/a	4.6	8.3	4.5
Track 2								
Main	-24.4	10.5	23.3	-1.9	-18.3	-19.2*	-2.8	-28.5***
Change in composition controls[a]	-18.6	17.3	30.7*	7.2	-14.2	-13.5	5.6	-20.6*
PY 1 to PY 3 change[b]	n/a	-0.2	6.8	2.5	n/a	-16.8	-17.6*	-25.8**

Source: Mathematica's analysis of Medicare claims data from January 2016 through December 2021; 2017 and 2018 CPC+ Practice Survey; and 2017 Q1–Q2 and 2018 Q3–Q4 CPC+ Practice Portal data.

Note: The follow-up, identification of patients for intervention, and same-day visit measures ranged from 0 to 1, with 1 being the highest score possible and indicating the practices used the process or offered the service to the fullest extent possible as measured by the variable(s) included in the measure. The LCM measure is on a scale of 0 to 100 as the percentage of patients enrolled in LCM across all risk tiers. The CMs per patient measure is in units of FTEs per 1,000 patients. Each row represents a separate regression model that included practice-level controls. Therefore, we generated these estimates from models that do not control for the other care process measures although they control for practice, geographic, and beneficiary characteristics at the practice level. These controls include HRR fixed effects to help control for effects COVID-19 may have had in PY 4 and PY 5. See methods section for list of controls.

[a] Coefficients from these rows are from regressions that additionally control for the change in the average beneficiary's HCC score, age, and dual-eligibility status at the practice from baseline to the relevant program year (indicated by the column).

[b] Coefficients from these rows are from regressions in which instead of using the practice's change in the care process measure from PY 1 to PY 2 as the regressor, we used the change from PY 1 to PY 3.

[c] Coefficients from these rows are from regressions in which instead of using the care process defined as CMs per 1,000 patients, we used CMs per 100 attributed Medicare beneficiaries.

*/**/*** Significantly different from zero at the 0.10/0.05/0.01 level, two-tailed test.

CM = care manager; Coef = coefficient; ED = emergency department; FTE = full-time equivalent; HCC = hierarchical condition category; HRR = hospital referral region; LCM = longitudinal care management; n/a = not applicable; SE = standard error; pp = percentage point; PY = Program Year.

B. Additional outcomes

To understand the mechanisms for the robust and favorable associations between same-day visits and outpatient ED visits and acute hospitalizations, we looked at the associations for more specific outcomes, which are subsets of the outpatient ED visit and acute hospitalization outcomes. For outpatient ED visits, we additionally assessed the effect on PCS ED visits and preventable ED visits (mutually exclusive categories of outpatient ED visits). For acute hospitalizations, we additionally looked at total *medical acute hospitalizations* and *medical acute hospitalizations without a CC or MCC,* which are a subset of *medical acute hospitalizations.*

For Track 1, we found that effects of changes in same-day visits on changes in the **AHR** were generally more concentrated in these more specific categories that are lower acuity and more likely preventable—shown by the fact that the percentage impacts were larger. We did not find this for the more specific categories of **outpatient ED visits** or for Track 2. These results provide some suggestive evidence that improving same-day visits leads to a **lower AHR** by primarily reducing lower acuity hospitalizations. Alternatively, improving same-day visits leads to fewer **outpatient ED visits** by reducing PCS ED visits, preventable visits, and other types of visits not included in those categories.

Table 5.O.5. Associations of changes in same-day visits and changes in additional outcomes over the CPC+ intervention period

	Baseline to PY 2		Baseline to PY 3		Baseline to PY 4		Baseline to PY 5	
	Effect	% effect	Effect	% effect	Effect	% effect	Effect	% effect
Track 1								
Outpatient ED visits	-42**	-8%	-35*	-7%	-46***	-9%	-27	-5%
Primary care substitutable ED visits	-18**	-9%	-12	-6%	-18**	-9%	-4	-2%
Preventable ED visits	-13**	-10%	-10	-8%	-14**	-11%	-5	-4%
Acute hospitalizations	-25**	-9%	-27**	-9%	-6	-2%	-14	-5%
Medical acute hospitalizations	-20**	-10%	-27***	-14%	-8	-4%	-11	-5%
Medical acute hospitalizations without a CC or MCC	-14***	-18%	-11***	-15%	-7*	-10%	-9**	-12%
Track 2								
Outpatient ED visits	-24	-5%	10	2%	23	5%	-2	0%
Primary care substitutable ED visits	-2	-1%	4	2%	11	6%	1	0%
Preventable ED visits	-12*	-9%	-1	-1%	5	4%	-4	-3%
Acute hospitalizations	-18	-6%	-19*	-7%	-3	-1%	-28***	-10%
Medical acute hospitalizations	-8	-4%	-15	-8%	-2	-1%	-20**	10%
Medical acute hospitalizations without a CC or MCC	0	0%	-3	-4%	2	2%	1	1%

Source: Mathematica's analysis of Medicare claims data from January 2016 through December 2021; 2017 and 2018 CPC+ Practice Survey; and 2017 Q1–Q2 and 2018 Q3–Q4 CPC+ Practice Portal data.

Note: The same-day visit measure ranged from 0 to 1, with 1 being the highest score possible and indicating the practices used the process or offered the service to the fullest extent possible as measured by the variable(s) included in the measure. Each row represents a separate regression model that included practice-level controls. Therefore, we generated these estimates from models that do not control for the other care process measures although they control for practice, geographic, and beneficiary characteristics at the practice level. These controls include HRR fixed effects to help control for effects COVID-19 may have had in PY 4 and PY 5. See methods section for list of controls.

*/**/*** Significantly different from zero at the 0.10/0.05/0.01 level, two-tailed test.

Table 5.O.5. ***(continued)***

CC = complication or comorbidity; CM = care manager; Coef = coefficient; ED = emergency department; HRR = hospital referral region; LCM = longitudinal care management; MCC = major complication or comorbidity; SE = standard error; pp = percentage point; PY = Program Year.

5.O.4. Discussion

We found robust, favorable results that improvements in availability of same-day visits were associated with decreases in outpatient ED visits and acute hospitalizations. We had initially hypothesized that improvements in same-day visits would result in fewer outpatient ED visits due to patients going to their primary care doctor for non-emergent and primary care treatable ailments. It is perhaps a bit surprising that we found effects on acute hospitalizations, where we expected that these visits could not have easily been substituted with primary care at the point when patients needed care. However, one explanation is that more availability of same-day and next-day visits encourages patients to go to the doctor *before* something becomes serious, resulting in an acute hospitalization. That is, without the availability of same-day visits, patients may just "wait and see" if their condition improves (rather than going through the hassle of the ED), and in some cases, this leads them to develop a much more serious condition, which results in hospitalization. Possible examples of conditions that might fit this theorized mechanism are urinary tract infections and congestive heart failure.

Although the same-day visit results were fairly robust, they still suffer from potential limitations. The largest is that, even in PY 1, most practices already reported having same-day visits available for all patients who needed them, so there was not much room for improvement over the intervention period. Our analysis, therefore, relied on a small amount of variation.

We interpret our results as suggesting the potential importance of availability of same-day or next-day visits and that future research is needed to cement the relationship. One research improvement that could be addressed in the future is creating survey measures that are more specific and have more room for improvement at baseline compared to the measures we used. For example, a measure of the *number* of the primary care visit slots that are reserved for same-day visits would likely have more specificity and variation across practices than our current measure of the rough portion of patients for which they have same-day and next-day visits available. In addition, the number or portion of primary care practitioners at a practice who provide same-day visits might help illuminate whether there is the potential for interpersonal continuity.

For the other four care process measures, the fact that we found inconsistent (and in some cases unfavorable) results does not necessarily mean the care process measures have no impact on outcomes. The LCM and CMs per patient measures were particularly noisy, driven partly by large changes in their denominator—the total number of patients—from PY 1 to PY 2. For LCM, there were also very large and unrealistic changes in the number of patients under LCM. The noisiness of the measures may have contributed to the null findings. In addition, changes in these measures might be driven by changes in the characteristics of patients (for example, if the patient population became riskier, or was anticipated to become riskier, practices might hire more care managers or increase the portion of patients under LCM). Although we tried to control for changes in the composition of the beneficiary sample in our sensitivity tests, we might not have been able to do this completely, resulting in a bias towards unfavorable findings.

Some of the promising effects we found in other analyses (reported in Appendices 5.L, 5.M, and 5.N) were not present in this analysis. That is partly because of differences in controls (such as the inclusion of HRR fixed effects for this analysis), and partly because of differences in outcomes. In Appendix 5.N. Exemplar Quantitative, we found promising evidence that the follow-up and LCM measures were

significantly negatively associated with changes in Bayesian-adjusted outpatient ED visits and acute hospitalizations from baseline to PY 2. The difference in results from our analysis and the Exemplar Quantitative analysis appears to be driven by a combination of lower standard errors using the Bayesian-adjusted outcomes and slightly more favorable estimates. The Bayesian-adjusted outcomes more flexibly controlled for risk of beneficiaries, which might have more completely adjusted for changes in composition of beneficiaries; this could help explain the more favorable Exemplar Quantitative estimates.

5.O.5. Conclusion

Based on our analysis, improvements in availability of same-day and next-day visits showed the strongest effects on improving outpatient ED visits and acute hospitalizations. Given the limitations of our study, further investigation is needed to better understand whether availability of same-day and next-day appointments truly affect ED visits and acute hospitalizations, and whether improvements in other care processes, where we found limited favorable associations, do not. To do so, researchers and model evaluators need better measures of these activities and accurate measurements of them over time, including measurements before the model starts. Survey items should be designed to measure care process activities explicitly and concretely, with ample room for growth. Using methods that more fully risk-adjust and account for noise in the outcome (like those used in the Appendix 5.N. Exemplar Quantitative) may make it easier to isolate impacts of care processes, even though these can be time-consuming to create. In addition to using better measures, future innovation models should explicitly test different strategies for changes in care delivery—for example, by randomly assigning care delivery process activities for participating practices to pursue. Such approaches could help policymakers understand the care processes that lead to improved outcomes, as well as the return on investments into these aspects of primary care.

5.P. Examining associations between practice capabilities and performance-based incentive payments

Key takeaways

We examined the associations between practice capabilities and CPC+ Performance-based Incentive Payments (PBIPs). We found that contemporaneous levels of many practice capabilities were positively associated with practices' PBIP quality scores. For Track 2 practices only, we found that changes (improvements) in practice capabilities were also positively associated with practices' PBIP quality scores. We found more limited evidence of associations between contemporaneous levels of and changes in practice capabilities and practices' PBIP utilization scores.

The positive associations between practice capabilities and PBIP scores found in this analysis are reassuring in that they show that the performance-based payments practices received in CPC+ were correlated with the practices' actions (capabilities) in the expected direction.

5.P.1. Introduction

In this Appendix, we first introduce the motivation and research question for the analysis (Section 1). We then explain the analytic methods (Section 2). Finally, we describe the results (Section 3) and discuss their implications (Section 4).

The health care landscape in the United States is rapidly shifting toward value-based care. One prominent feature of this transformation is the use of performance-based payments to reward providers for delivering high-quality and cost-efficient care (Chee et al. 2016). Performance-based systems like CPC+ PBIPs reward practices for achieving favorable health care outcomes for their beneficiaries. Practices cannot directly control these health care outcomes but can control their actions that influence these health care outcomes. Practice capability measures are a proxy for practices' actions. Practices' actions ideally should correlate with the rewards they receive under performance-based systems. The rich set of CPC+ and supplemental data available for the evaluation of CPC+ provides a unique opportunity to study the associations between practice capabilities and the performance-based payments they receive in the context of effectiveness of the CPC+ performance-based payment system.

In a previous analysis, we found that higher CPC+ PBIPs were associated with not only lower levels of service use outcomes (as designed) but also improvements in these outcomes relative to the baseline. CPC+ PBIPs were also positively associated with both lower levels and lower growth in expenditures even though expenditures were not part of the PBIP calculations (Peikes et al. 2021b). Although not causal, the findings suggested that a performance-based payment methodology focused on attaining benchmark performance levels (like CPC+ PBIPs) could reward practices that improve beyond the benchmark level or reduce expenditures, without involving more complicated incentives explicitly related to expenditures or continuous improvement.

In our current analysis, we examine the associations between contemporaneous levels of and changes in practice capabilities and PBIPs in the first three program years (2017 through 2019). The objective of the analysis is to understand whether and how performance-based payments practices received correlated with their actions (capabilities).

5.P.2. Methods

A. Setting: PBIP methodology and components

At the beginning of each CPC+ program year, CMS prospectively paid the maximum amount of PBIP that practices were eligible to receive in that year (that is, $2.50 per beneficiary per month [PBPM] for Track 1 and $4.00 PBPM for Track 2 multiplied by the number of eligible Medicare fee-for-service [FFS] beneficiaries attributed to the practice). In the year following the performance year, CMS recouped PBIPs based on a total PBIP score ranging from 0 to 1 that was based on the average of the quality and utilization components. For example, a practice with a total PBIP score of 0.50 retained 50 percent of the total PBIP.

Informed by behavioral economics theory, PBIPs were designed to test whether timely payments (via prospective, maximum payments) combined with loss aversion (to avoid retrospective recoupments) provide practices with greater incentives to achieve the goals of CPC+, as compared to a conventional retrospective performance-based payment approach, where payment is not made until well after the end of each performance year (Audet and Zezza 2015; Khuller et al. 2015). CMS shared measure performance results and final PBIP scores with practices so they could identify areas for improvement.

CMS calculated the PBIP quality score using eCQMs (constituting 75 percent of the score in 2017 and 2018 and 60 percent of the score in 2019) and patient experience-of-care measures from the CG CAHPS Patient-Centered Medical Home Survey (constituting the remaining 25 percent of the score in 2017 and 2018 and 40 percent of the score in 2019). CMS calculated the PBIP utilization score using claims-based inpatient hospital utilization (constituting 67 percent of the score) and emergency department (ED) utilization (constituting the remaining 33 percent of the score). Higher PBIP quality scores and PBIP utilization scores indicated better performance (that is, higher quality and lower utilization). Each component score ranged from 0 to 1, based on the practice's performance relative to benchmarks. Practices that did not meet the minimum reporting and performance requirements in the quality component were not eligible to retain any PBIP regardless of how they performed on utilization.

The PBIP design was simple to implement and allowed CMS to establish clear performance goals and to provide timely feedback to practices regarding their performance. The continuous scale reflected the intention to reward practices for every increment by which they outperformed the benchmark.

The PBIP design generally did not explicitly incentivize high levels or improvements in the measures of practice capabilities used in this analysis (see Table 5.P.1 below). The sole exceptions were the three practice capability measures in the Improve Access domain, which are directly related to the patient experience-of-care measures included in the PBIP quality domain. However, even without explicit incentives, it is possible that the PBIPs could have indirectly encouraged practices to improve on practice capabilities as they sought to achieve higher scores.

Practices participating in both CPC+ and a Medicare Shared Savings Program (SSP) Accountable Care Organization (ACO) were not eligible to receive PBIPs because they were eligible to receive a portion of any shared savings earned by their ACO, and CMS rules prohibited "double dipping."

B. Study sample

Our study sample included primary care practices that participated in CPC+ across 14 regions from January 2017 through December 2019. In response to the coronavirus disease 2019 (COVID-19) public health emergency, CMS modified the PBIP methodology for 2020; therefore, we excluded 2020 data from our analysis. CMS reverted to the original PBIP methodology in 2021; however, we did not receive 2021 PBIP data in time to include 2021 PBIP scores in our analysis.

For each program year, we restricted the sample to PBIP-eligible practices (that is, those that did not participate in Medicare SSP at the beginning of the year). As practices joined or left SSP, the sample of practices eligible for PBIPs changed. This resulted in 1,763 unique practices across the three years. These practices collectively provided care for 1.21 million Medicare FFS beneficiaries in 2016—the year before CPC+ began (baseline year). Our sample included 1,518 practices in 2017 (627 in Track 1 and 891 in Track 2), 1,371 practices in 2018 (548 in Track 1 and in 823 Track 2), and 1,430 practices in 2019 (566 in Track 1 and 864 in Track 2). Of the 1,763 practices, 1,184 (67.2 percent) remained as non-SSP participants in all three years.

C. Data sources

We assembled a practice-year dataset that included four types of data: (1) practice characteristics at the start of CPC+ (2016); (2) practices' PBIP scores (quality and utilization scores) from 2017 through 2019; (3) characteristics of beneficiaries, aggregated to the practice level from 2017 through 2019; and (4) practices' responses on survey questions related to practice capabilities from 2017 through 2019. See Table 5.P.1 for all data sources.

Table 5.P.1. Variables and data sources

Variable	Data source
Practice capabilities	
Outreach after emergency department visits and hospitalizations	2017, 2018, and 2019 CPC+ Practice Surveys
Percentage of patients targeted for care management	2017, 2018, and 2019 CPC+ Practice Surveys
Pre-visit planning (gathering patient info)	2017, 2018, and 2019 CPC+ Practice Surveys
Stratifying patients by risk level	2017, 2018, and 2019 CPC+ Practice Surveys
Self-management support for patients	2017, 2018, and 2019 CPC+ Practice Surveys
Specific physician for patients	2017, 2018, and 2019 CPC+ Practice Surveys
Extended hours of operation including weekend	2017, 2018, and 2019 practice-reported care delivery data submitted to CMS
24/7 patient access to care team	2017, 2018, and 2019 practice-reported care delivery data submitted to CMS
Same- or next-day access to appointments	2017, 2018, and 2019 practice-reported care delivery data submitted to CMS
Have QI specialist	2017, 2018, and 2019 CPC+ Practice Surveys
Organize and support QI activities	2017, 2018, and 2019 CPC+ Practice Surveys
Staff, resources, and time for QI activities	2017, 2018, and 2019 CPC+ Practice Surveys
Use feedback from patient surveys	2017, 2018, and 2019 CPC+ Practice Surveys
Use registry to identify and track patients	2017, 2018, and 2019 CPC+ Practice Surveys
Practice characteristics	
Number of practitioners with primary care specialty	SK&A 2016, NPPES 2016
Whether practice is owned by a hospital or health system	SK&A 2016
Whether practice is multispecialty	SK&A 2016
Meaningful use status (whether physicians at practice had attested to meaningful use of EHRs and earliest year that physician at practice became meaningful user)	CMS 2016
Whether in an urban, rural, or suburban area	Area Resource File, 2015–2016
Medicare price index of the hospital referral region	CMS' Medicare Geographic Variation data, 2015
Characteristics of Medicare beneficiaries attributed to practices	
Demographic mix of attributed beneficiaries (percentage of practice in age, race, and gender categories)	Medicare enrollment data, 2014–2019
Distribution of Medicare risk scores (HCC)	2015–2018 risk scores computed from Medicare claims and enrollment data
Percentage of practice's attributed Medicare beneficiaries with 22 chronic conditions defined by HCCs or CCW algorithm	Medicare claims data, 2013–2019
Percentage of adults age 25 or older in the county with a degree from a four-year college	Area Resource File, 2017
Percentage in categories for original reason for Medicare entitlement	Medicare enrollment data, 2014–2019
Percentage of beneficiaries dually eligible for Medicaid	Medicare enrollment data, 2014–2019
Median household income of county	Area Resource File, 2017

Note: We defined practice characteristics at the start of CPC+. Beneficiary characteristics are defined in each program year, except for two: (1) percentage of adults age 25 or older in the county with a degree from a four-year college and (2) median household income of county. These variables are defined at the county level of the practice, but we use them as proxies for beneficiaries' education and income levels.

CCW = Chronic Conditions Warehouse; CMS = Centers for Medicare & Medicaid Services; ED = emergency department; EHR = electronic health record; HCC = Hierarchical Condition Category; NPPES = National Plan & Provider Enumeration System; QI = quality improvement.

Practice characteristics. We began with a practice-level dataset constructed for the independent evaluation of CPC+ (Peikes et al. 2021b). This dataset includes practice characteristics defined for the baseline year—the year before the start of CPC+ (2016). We kept three practice characteristics that were included in the SK&A data, specifically, counts of primary care practitioners (PCPs), ownership status (hospital or system owned/independent), and whether the practice is multispecialty.

To this dataset, we added other practice and market characteristics, such as urban/rural status and county median household income, using publicly available data (such as the Area Resource File), CMS restricted-use data (such as the Master Data Management), and proprietary data (such as National Committee for Quality Assurance data).

We added data on the characteristics of the practice's Medicare FFS beneficiaries using Medicare claims and enrollment data from 2014 through 2019. We assigned Medicare beneficiaries to practices where they had a chronic care management visit or an Annual Wellness Visit, or where they received the largest number of primary care visits (see Appendix 6.B of Ghosh et al. [2020] for more details on the assignment process). For each practice, we calculated the proportion of beneficiaries in each age group (under 65, 65 to 74, 75 to 84, and 85 or older), sex (male), race (White, Black, all other races), original reason for Medicare entitlement (old age, disability, or end-stage renal disease (ESRD)/ESRD and disability combined), dual eligibility for Medicaid, chronic conditions (captured by Hierarchical Condition Category [HCC] and Chronic Conditions Warehouse [CCW] algorithm) and the mean HCC risk score of beneficiaries assigned to each practice.

PBIP scores. We used the final quality and utilization PBIP scores for 2017, 2018, and 2019 calculated by CMS's CPC+ payment contractor.

Practice capabilities. We used practice-reported measures of practice capabilities from the CPC+ practice surveys we fielded in 2017, 2018, and 2019 and from practice-reported care delivery data submitted to CMS (which we refer to as CDR [care delivery requirement] data). The surveys asked CPC+ practices to report the degree to which they implemented various practice capabilities.

D. Analysis

Selecting practice capabilities. We chose to include measures of practice capabilities in our analysis that met the following criteria:

1. Were measured consistently in 2017, 2018, and 2019
2. Had a potential connection to the measures of utilization or quality used to calculate PBIP utilization and quality scores
3. Were not topped out and had sufficient variation across practices and across years

Table 5.P.2 lists the practice capabilities we selected for this analysis.

Table 5.P.2. Practice capabilities included in PBIP analysis

Practice capability	Category	Scale	Included in PBIP quality analysis	Included in PBIP utilization analysis
Outreach after emergency department visits and hospitalizations	Expand care management	0 to 1	No	Yes
Percentage of patients targeted for care management	Expand care management	0 to 100	Yes	Yes
Pre-visit planning (gathering patient info)	Expand care management	0 to 1	Yes	No
Stratifying patients by risk level	Expand care management	0 to 1	Yes	No
Self-management support for patients	Expand care management	0 to 1	Yes	No
Specific physician for patients	Continuity of care	0 to 1	Yes	No
Extended hours of operation including weekend	Improve access	0 to 1	Yes	Yes
24/7 patient access to care team	Improve access	0 to 1	Yes	Yes
Same- or next-day access to appointments	Improve access	0 to 1	Yes	Yes
Have QI specialist	Invest in transformation	0 to 1	Yes	No
Organize and support QI activities	Invest in transformation	0 to 1	Yes	No
Staff, resources, and time for QI activities	Invest in transformation	0 to 1	Yes	No
Use feedback from patient surveys	Use of data	0 to 1	Yes	No
Use registry to identify and track patients	Use of data	0 to 1	Yes	No

Source: Mathematica's analysis of practice-reported care delivery data submitted to CMS and CPC+ Practice Survey data from 2017 through 2019.

PBIP = Performance-based Incentive Payment; SSP = Medicare Shared Savings Program, QI = quality improvement.

The percentage of patients targeted for care management measure ranges from 0 to 100 and is calculated across all risk tiers. All other practice capability measures included in our analysis are categorical. We converted the categorical practice capability measures to a scale of 0 to 1, with 0 corresponding to not displaying the practice capability at all and 1 corresponding to fully displaying the practice capability. Table 5.P.3 provides examples of how we rescaled these categorical variables.

Table 5.P.3. Rescaling practice capability measures

Example survey question type	Survey response	Rescaled value
How much of Activity A is at your practice?	A lot	1.00
	Some	0.50
	None	0.00
How often is Activity B at your practice?	Always	1.00
	Often	0.66
	Rarely	0.33
	Never	0.00

Multivariate regression to examine associations. We estimated two sets of multivariate linear regression models to understand the association between practices' contemporaneous levels of and changes in practice capabilities and PBIP quality scores (set one) and PBIP utilization scores (set two).

For the first set of regressions, we estimated a separate regression model for each practice capability where practices' PBIP scores in a particular year are explained by the contemporaneous level of the given practice capability. Each contemporaneous level regression model included observations from 2017, 2018, and 2019.

For the second set of regressions, we estimated a separate regression model for each practice capability where practices' PBIP scores in a particular year are explained by changes in the practice capability between 2017 and the current year, controlling for the 2017 level of the practice capability. Each change since the 2017 regression model included observations from 2018 and 2019.

We conducted the analyses separately by track and type of PBIP score (that is, quality and utilization).

Weighting to assess relationships by practice and by beneficiaries. The relationships between a practice's performance (PBIP scores) and its practice capabilities examined as part of this analysis could be used to inform and guide payment system design in future payment reform models. Therefore, we weighted the performance of each practice equally (regardless of its size) in assessing these relationships. To account for correlation in scores within practices over time, we used cluster-robust standard errors, clustering at the practice level. We show statistical significance at the 0.01, 0.05, and 0.10 levels. All statistical analyses were conducted using Stata, version 15.1 (StataCorp, LLC).

Control variables. All regression models include baseline practice characteristic, baseline beneficiary characteristic, and current year beneficiary characteristic control variables listed in Table 5.P.4. We included control variables for not only baseline but also current year beneficiary characteristics, to account for changes in beneficiary composition from 2017 to 2019.

Table 5.P.4. Control variables

Variable	Variable version(s)
Practice characteristics	
Number of practitioners with primary care specialty	Baseline
Whether practice is owned by a hospital or health system	Baseline
Whether practice is multispecialty	Baseline
Meaningful use status (whether physicians at practice had attested to meaningful use of EHRs and earliest year that physician at practice became meaningful user)	Baseline
Whether in an urban, rural, or suburban area	Baseline
Medicare price index of the hospital referral region	Baseline
Characteristics of Medicare beneficiaries attributed to practices	
Demographic mix of attributed beneficiaries (percentage of practice in age, race, and gender categories)	Baseline, current year
Distribution of Medicare risk scores (HCC)	Baseline, current year
Percentage of practice's attributed Medicare beneficiaries with 22 chronic conditions defined by HCCs or CCW algorithm)	Baseline, current year
Percentage of adults age 25 or older in the county with a degree from a four-year college	Baseline, current year
Percentage in categories for original reason for Medicare entitlement	Baseline, current year
Percentage of beneficiaries dually eligible for Medicaid	Baseline, current year
Median household income of county	Baseline, current year

Note: We defined practice characteristics at the start of CPC+. Current year beneficiary characteristics are defined in each program year, except for two: (1) percentage of adults age 25 or older in the county with a degree from a four-year college and (2) median household income of county. These variables are defined at the county level of the practice, but we use them as proxies for beneficiaries' education and income levels.

CCW = Chronic Conditions Warehouse; ED = emergency department; EHR = electronic health record; HCC = Hierarchical Condition Category.

5.P.3. Results

A. Summary statistics

PBIP scores were higher for quality than for utilization and grew over time for both components of the payment. In 2019 (the final year of our study), the average Track 1 CPC+ practice retained 84 percent ($1.05 PBPM) of its quality component and 52 percent ($0.65 PBPM) of its utilization component; the corresponding numbers for Track 2 practices were 86 percent ($1.72 PBPM) and 59 percent ($1.18 PBPM) for the quality and utilization components, respectively (Table 5.P.5). In both tracks, the mean PBIP scores were higher for both utilization and quality components in 2018 and 2019 relative to 2017 (the first year of the intervention).

Table 5.P.5. Summary of PBIP scores in the first three program years of CPC+

	Track 1			Track 2		
	Percentage of practices that did not retain any PBIP	Percentage of practices that retained full PBIP	Mean score (IQR)	Percentage of practices that did not retain any PBIP	Percentage of practices that retained full PBIP	Mean score (IQR)
Quality PBIP						
2017	4.6	16.9	0.72 (0.61, 0.89)	0.7	18.9	0.76 (0.65, 0.89)
2018	1.3	51.6	0.85 (0.73, 1.00)	0.0	56.1	0.89 (0.74, 1.00)
2019	1.1	56.7	0.84 (0.60, 1.00)	0.5	60.9	0.86 (0.60, 1.00)
Utilization PBIP						
2017	35.3	4.2	0.37 (0.00, 0.67)	26.4	3.4	0.43 (0.00, 0.70)
2018	19.5	4.7	0.51 (0.22, 0.78)	16.5	4.7	0.52 (0.24, 0.83)
2019	20.1	5.7	0.52 (0.22, 0.80)	12.2	6.3	0.59 (0.33, 0.87)

Source: Mathematica's analysis of PBIP performance data from 2017 through 2019.

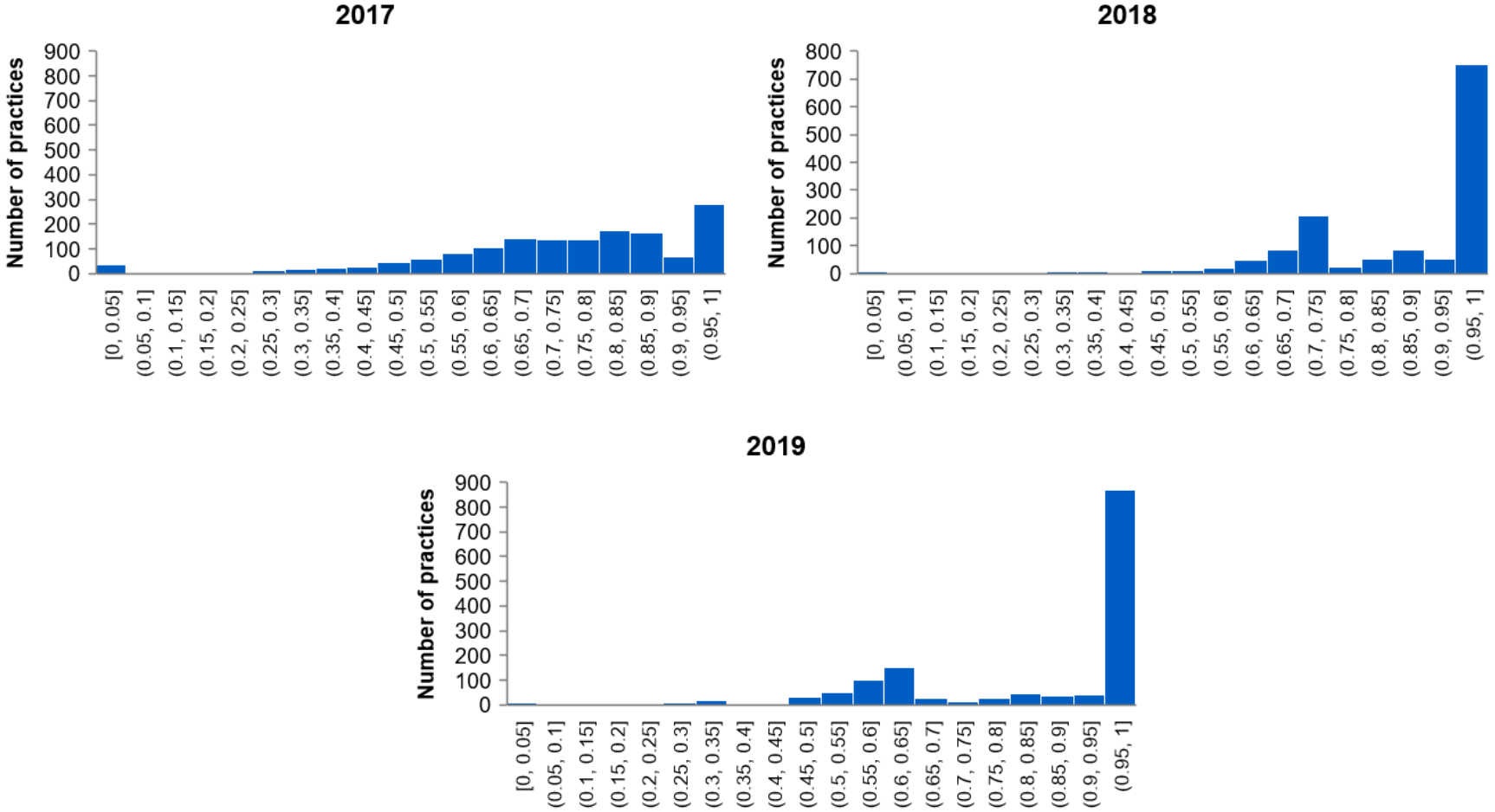
Note: Each PBIP score ranges from 0 to 1, representing the proportion of the total PBIP amount practices retained in each program year. Higher PBIP quality scores and PBIP utilization scores indicate better performance (that is, higher quality and lower utilization). Analysis includes CPC+ practices that did not participate in Medicare SSP in each program year of CPC+ because only non-SSP practices were eligible to receive the PBIP. SSP status can change from year to year as practices join and exit the program. For Track 1, we included 627, 548, and 566 practices in 2017, 2018, and 2019, respectively. For Track 2, we included 891, 823, and 864 practices in 2017, 2018, and 2019, respectively. All practices were weighted equally, irrespective of their size.

IQR = Inter-quartile range; PBIP = Performance-based Incentive Payment; SSP = Medicare Shared Savings Program.

Figure 5.P.1 illustrates PBIP quality scores were widely distributed across practices in 2017 and consolidated toward a majority of practices having perfect scores of 1 in 2018 and 2019. Many practices received PBIP utilization scores of 0 for performing poorly on utilization measures or not meeting the quality component requirements, especially in 2017 (Figure 5.P.2).[163] Otherwise, PBIP utilization scores were fairly evenly distributed between 0.15 and 1.0 in 2017, 2018, and 2019.

[163] Among practices receiving a PBIP utilization score of 0, roughly one-fifth (19 percent in 2017, 13 percent in 2018, and 22 percent in 2019) received a PBIP utilization score of 0 for not meeting the minimum quality reporting requirements.

Figure 5.P.1. Distribution of PBIP quality scores in the first three program years of CPC+

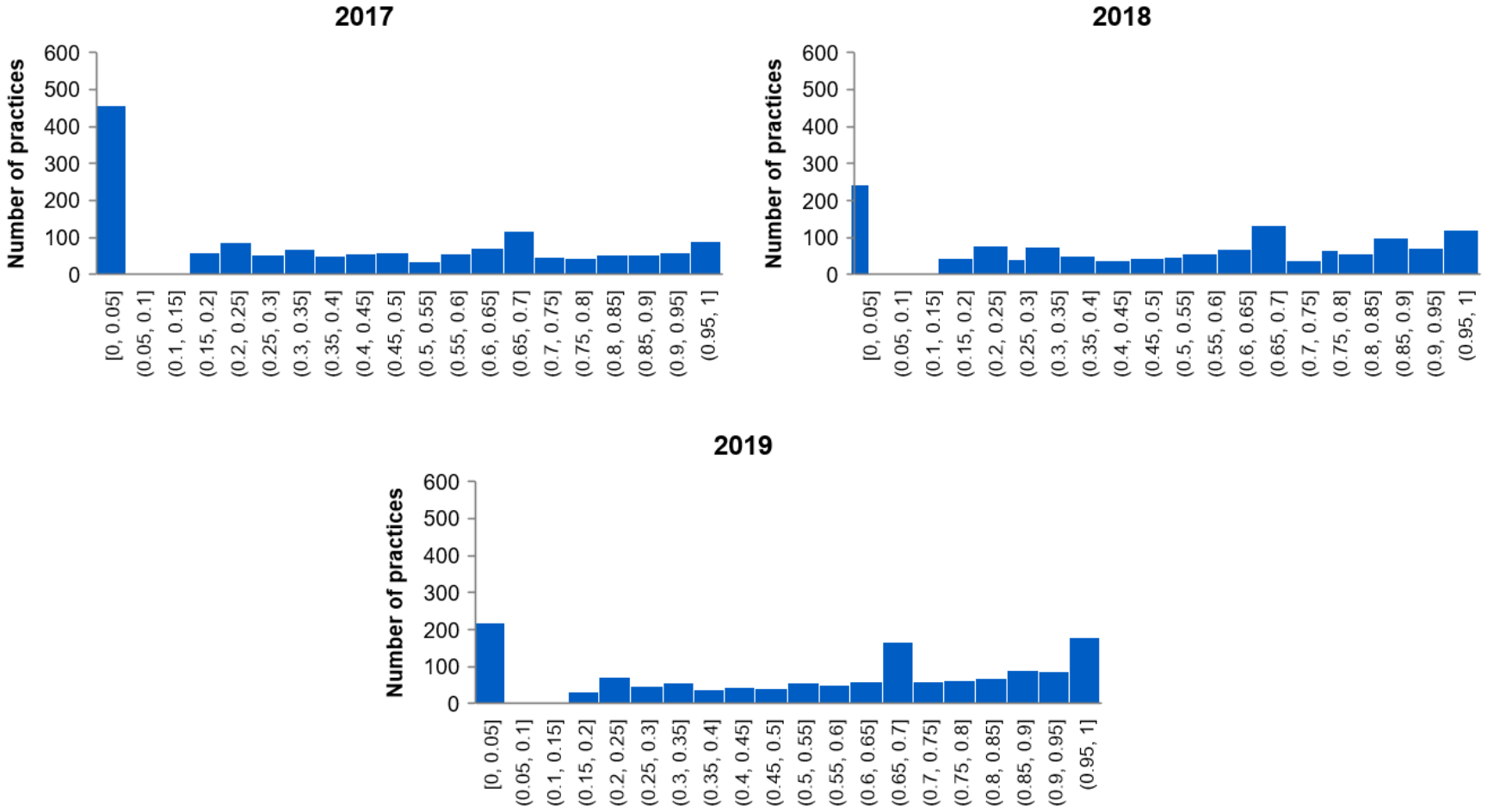


Source: Mathematica's analysis of PBIP performance data from 2017 through 2019.

Note: Each PBIP score ranges from 0 to 1, representing the proportion of the total PBIP amount practices retained in each program year. Higher PBIP quality scores and PBIP utilization scores indicate better performance (that is, higher quality and lower utilization). Analysis includes CPC+ practices that did not participate in Medicare SSP in each program year of CPC+ because only non-SSP practices were eligible to receive the PBIP. SSP status can change from year to year as practices join and exit the program. We included 1,518, 1,371, and 1,430 practices in 2017, 2018, and 2019, respectively. All practices were weighted equally, irrespective of their size.

PBIP = Performance-based Incentive Payment; SSP = Medicare Shared Savings Program.

Figure 5.P.2. Distribution of PBIP utilization scores in the first three program years of CPC+



Source: Mathematica's analysis of PBIP performance data from 2017 through 2019.

Note: Each PBIP score ranges from 0 to 1, representing the proportion of the total PBIP amount practices retained in each program year. Higher PBIP quality scores and PBIP utilization scores indicate better performance (that is, higher quality and lower utilization). Analysis includes CPC+ practices that did not participate in Medicare SSP in each program year of CPC+ because only non-SSP practices were eligible to receive the PBIP. SSP status can change from year to year as practices join and exit the program. We included 1,518, 1,371, and 1,430 practices in 2017, 2018, and 2019, respectively. All practices were weighted equally, irrespective of their size.

PBIP = Performance-based Incentive Payment; SSP = Medicare Shared Savings Program.

For almost all the practice capability measures, the mean across practices increases from 2017 to 2019 (Table 5.P.6), indicating that practices on average more fully implemented these practice capabilities over time. For the few practice capability measures with little change over time or means close to or above 0.90, there was still substantial within-practice change, with at least one-third of practices reporting a different level of each practice capability in 2019 compared with 2017.

Table 5.P.6. Degree of implementation of practice capabilities included in PBIP analysis

Practice capability measure	Category	Scale	2017 mean	2018 mean	2019 mean	Percentage of practices with change (2017 to 2019)
Outreach after emergency department visits and hospitalizations	Expand care management	0 to 1	0.77	0.89	0.93	54.5%
Percentage of patients targeted for care management	Expand care management	0 to 100	6.25	7.23	5.78	100%
Pre-visit planning (gathering patient info)	Expand care management	0 to 1	0.60	0.69	0.73	57.9%
Stratifying patients by risk level	Expand care management	0 to 1	0.58	0.82	0.83	63.9%
Self-management support for patients	Expand care management	0 to 1	0.56	0.70	0.76	64.7%
Specific physician for patients	Continuity of care	0 to 1	0.87	0.89	0.88	34.8%
Extended hours of operation including weekend	Improve access	0 to 1	0.70	0.75	0.78	40.7%
24/7 patient access to care team	Improve access	0 to 1	0.84	0.89	0.90	52.9%
Same- or next-day access to appointments	Improve access	0 to 1	0.91	0.92	0.93	41.5%
Have QI specialist	Invest in transformation	0 to 1	0.38	0.45	0.49	40.3%
Organize and support QI activities	Invest in transformation	0 to 1	0.75	0.84	0.88	50.5%
Staff, resources, and time for QI activities	Invest in transformation	0 to 1	0.58	0.67	0.72	58.1%
Use feedback from patient surveys	Use of data	0 to 1	0.72	0.84	0.87	57.7%
Use registry to identify and track patients	Use of data	0 to 1	0.63	0.66	0.75	52.4%

Source: Mathematica's analysis of practice-reported care delivery data submitted to CMS and CPC+ Practice Surveys data from 2017 through 2019.

Note: For practice capability measures, higher values indicate a higher degree of implementation of the practice capability. The percentage of patients targeted for care management measure is on a scale of 0 to 100 as the percentage of patients enrolled across all risk tiers. Analysis includes CPC+ practices that did not participate in Medicare SSP in each program year of CPC+ because only non-SSP practices were eligible to receive the PBIP. SSP status can change from year to year as practices join and exit the program. We included 1,518, 1,371, and 1,430 practices in 2017, 2018, and 2019, respectively. All practices were weighted equally, irrespective of their size.

PBIP = Performance-based Incentive Payment; SSP = Medicare Shared Savings Program; QI = quality improvement.

B. Associations between PBIPs and contemporaneous levels of practice capabilities

For a majority of the practice capabilities included in our analysis, we found positive associations between contemporaneous levels of the practice capability and PBIP quality scores. We found strong consistency between Track 1 and Track 2 practices; for 8 of the 13 practice capabilities, contemporaneous levels of the practice capability had a statistically significant positive association with PBIP quality scores for both Track 1 practices and Track 2 practices (Table 5.P.7). For example, the difference between a practice not stratifying patients by risk level and a practice fully stratifying patients by risk level (that is, going from a

value of 0 to 1) was associated with a 0.10 higher PBIP quality score for Track 1 practices and 0.09 for Track 2 practices.

Contemporaneous levels of the three practice capabilities in the improve access category (extended hours of operation including weekend, 24/7 patient access to care team, and same- or next-day appointments) showed much weaker associations with PBIP quality scores than practice capabilities in the other categories. Same- or next-day access to appointments was the only practice capability in the improve access category for which the contemporaneous level had a statistically significant positive association with PBIP quality scores and that was for Track 2 practices only. In contrast, contemporaneous levels of practice capabilities had a statistically significant positive association with PBIP quality scores for 8 of the 10 practice capabilities in the other categories (expand care management, continuity of care, invest in transformation, and use of data) for both Track 1 and Track 2 practices.

Table 5.P.7. Regression results on the correlation between practices' yearly PBIP quality scores and contemporaneous levels of practice capabilities, from 2017 to 2019

		Quality PBIP	
Practice capabilities	Category	Track 1 Impact estimate (SE)	Track 2 Impact estimate (SE)
Percentage of patients targeted for care management	Expand care management	-0.0002 (0.0003)	0.0001 (0.0004)
Pre-visit planning (gathering patient info)	Expand care management	***0.05** (0.02)***	***0.08*** (0.02)***
Stratifying patients by risk level	Expand care management	***0.10*** (0.02)***	***0.09*** (0.02)***
Self-management support for patients	Expand care management	***0.05*** (0.02)***	***0.04*** (0.01)***
Specific physician for patients	Continuity of care	***0.12*** (0.04)***	***0.07*** (0.03)***
Extended hours of operation including weekend	Improve access	0.00 (0.02)	0.00 (0.01)
24/7 patient access to care team	Improve access	0.05 (0.04)	0.03 (0.02)
Same- or next-day access to appointments	Improve access	0.06 (0.05)	***0.15*** (0.04)***
Have QI specialist	Invest in transformation	0.01 (0.01)	0.01 (0.01)
Organize and support QI activities	Invest in transformation	***0.08*** (0.02)***	***0.04** (0.02)***
Staff, resources, and time for QI activities	Invest in transformation	***0.06** (0.02)***	***0.05*** (0.01)***
Use feedback from patient surveys	Use of data	***0.04** (0.02)***	***0.06*** (0.02)***
Use registry to identify and track patients	Use of data	***0.04** (0.02)***	***0.03** (0.01)***

Sources: Mathematica's analysis of practice-reported care delivery data submitted to CMS, CPC+ Practice Survey data, and PBIP performance data from 2017 through 2019 and Medicare Enrollment Database from January 2014 through December 2019.

Note: This table shows the results of 26 regressions, one for each practice capability by track. In each regression, we controlled for baseline practice characteristics, baseline beneficiary characteristics, and current year beneficiary characteristics. All practices were weighted equally, irrespective of their size. Each PBIP score ranges from 0 to 1. Higher PBIP quality scores and PBIP utilization scores indicate better performance (that is, higher quality and lower utilization). The analysis included only non-SSP practices because SSP practices were not eligible to receive PBIPs. ***Yellow shading* with bold, italicized text** signifies that the underlying coefficient was statistically significant at the 10 percent level or better using a two-sided test.

*** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$.

ED = emergency department; PBIP = Performance-based Incentive Payment; QI = quality improvement; SE = standard error; SSP = Medicare Shared Savings Program.

Contemporaneous levels of practice capabilities had more limited associations with PBIP utilization scores than PBIP quality scores. For Track 1 practices, contemporaneous levels of three of the five practice capabilities examined (percent of patients targeted for care management, outreach after ED visits and hospitalizations, and same- or next-day access to appointments) were positively and statistically significantly associated with PBIP quality scores (Table 5.P.8). In addition, somewhat counterintuitively, contemporaneous levels of 24/7 patient access to care team was negatively and statistically significantly associated with PBIP utilization scores for Track 1 practices. For Track 2 practices, outreach after ED visits and hospitalizations was the only practice capability for which contemporaneous levels were positively and statistically significantly associated with PBIP utilization scores.

Table 5.P.8. Regression results on the correlation between practices' yearly PBIP utilization scores and contemporaneous levels of practice capabilities, from 2017 to 2018 and 2018 to 2019

		Utilization PBIP	
Practice capabilities	Category	Track 1 Impact estimate (SE)	Track 2 Impact estimate (SE)
Percentage of patients targeted for care management	Expand care management	***0.0014*** (0.0005)***	0.0002 (0.0005)
Outreach after ED visits and hospitalizations	Expand care management	***0.11*** (0.03)***	***0.07** (0.03)***
24/7 patient access to care team	Improve access	***-0.07** (0.04)***	0.03 (0.04)
Same- or next-day access to appointments	Improve access	***0.11** (0.06)***	0.05 (0.05)
Extended hours of operation including weekend	Improve access	-0.01 (0.02)	-0.01 (0.02)

Sources: Mathematica's analysis of practice-reported care delivery data submitted to CMS, CPC+ Practice Survey data, and PBIP performance data from 2017 through 2019 and Medicare Enrollment Database from January 2014 through December 2019.

Note: This table shows the results of 10 regressions, one for each practice capability by track. In each regression, we controlled for baseline practice characteristics, baseline beneficiary characteristics, and current year beneficiary characteristics. All practices were weighted equally, irrespective of their size. Each PBIP score ranges from 0 to 1. Higher PBIP quality scores and PBIP utilization scores indicate better performance (that is, higher quality and lower utilization). The analysis included only non-SSP practices because SSP practices were not eligible to receive PBIPs. ***Yellow shading with bold, italicized text*** signifies that the underlying coefficient was statistically significant at the 10 percent level or better using a two-sided test.

*** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$.

ED = emergency department; PBIP = Performance-based Incentive Payment; QI = quality improvement; SE = standard error; SSP = Medicare Shared Savings Program.

C. Associations between changes in practice capabilities and PBIP scores

Changes in practice capabilities showed a similar pattern of associations with PBIP quality scores as contemporaneous levels for Track 2 practices but not for Track 1 practices. For Track 2 practices, changes in practice capabilities since 2017 were positively and statistically significantly associated with PBIP quality scores for seven of the nine practice capabilities for which contemporaneous levels were positively and statistically significantly associated with PBIP quality scores (Table 5.P.9). For example, going from not stratifying patients by risk level in 2017 to fully stratifying patients by risk level in 2018 or 2019 was associated with a 0.07 increase in 2018 or 2019 PBIP quality scores for Track 2 practices. For Track 1 practices, self-management support for patients was the only practice capability for which changes since 2017 were positively and statistically significantly associated with PBIP quality scores. For Track 1 practices, changes in the percentage of patients targeted for care management were negatively and statistically significantly associated with PBIP quality scores.

Table 5.P.9. Regression results on the correlation between practices' yearly PBIP quality scores and changes in practice capabilities, from 2017 to 2019

Practice capabilities	Category	Quality PBIP	
		Track 1 Impact estimate (SE)	Track 2 Impact estimate (SE)
Percentage of patients targeted for care management	Expand care management	***-0.0008**** ***(0.0006)***	-0.0007 (0.0006)
Pre-visit planning (gathering patient info)	Expand care management	0.00 (0.03)	***0.07**** ***(0.02)***
Stratifying patients by risk level	Expand care management	-0.01 (0.03)	***0.07**** ***(0.03)***
Self-management support for patients	Expand care management	***0.04**** ***(0.03)***	0.02 (0.02)
Specific physician for patients	Continuity of care	0.03 (0.06)	***0.14**** ***(0.04)***
Extended hours of operation including weekend	Improve access	-0.03 (0.02)	-0.02 (0.02)
24/7 patient access to care team	Improve access	0.00 (0.04)	0.03 (0.04)
Same- or next-day access to appointments	Improve access	-0.02 (0.06)	***0.12**** ***(0.06)***
Have QI specialist	Invest in transformation	0.02 (0.02)	0.00 (0.01)
Organize and support QI activities	Invest in transformation	0.02 (0.03)	***0.04**** ***(0.03)***
Staff, resources, and time for QI activities	Invest in transformation	-0.02 (0.03)	***0.07**** ***(0.02)***
Use feedback from patient surveys	Use of data	0.00 (0.04)	0.04 (0.03)
Use registry to identify and track patients	Use of data	0.03 (0.02)	***0.03**** ***(0.02)***

Sources: Mathematica's analysis of practice-reported care delivery data submitted to CMS, CPC+ Practice Survey data, and PBIP performance data from 2017 through 2019 and Medicare Enrollment Database from January 2014 through December 2019.

Note: This table shows the results of 26 regressions, one for each practice capability by track. In each regression, we controlled for baseline practice characteristics, baseline beneficiary characteristics, and current year beneficiary characteristics. All practices were weighted equally, irrespective of their size. Each PBIP score ranges from 0 to 1. Higher PBIP quality scores and PBIP utilization scores indicate better performance (that is, higher quality and lower utilization). The analysis included only non-SSP practices because SSP practices were not eligible to receive PBIPs. ***Yellow shading* with bold, italicized text** signifies that the underlying coefficient was statistically significant at the 10 percent level or better using a two-sided test.

*** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$.

ED = emergency department; PBIP = Performance-based Incentive Payment; QI = quality improvement; SE = standard error; SSP = Medicare Shared Savings Program.

Changes in practice capabilities showed a similar pattern of associations with PBIP utilization scores as contemporaneous levels of practice capabilities. For the practice capabilities for which contemporaneous levels were statistically significantly associated with PBIP utilization scores, there was also a statistically significant association of a similar magnitude between changes in the practice capability and the PBIP utilization scores, the only exception being the percentage of patients targeted for care management for Track 1 practices (Table 5.P.10).

Table 5.P.10. Regression results on the correlation between practices' yearly PBIP utilization scores and changes in practice capabilities, from 2017 to 2018 and 2018 to 2019

		Utilization PBIP	
Practice capabilities	Category	Track 1 Impact estimate (SE)	Track 2 Impact estimate (SE)
Percentage of patients targeted for care management	Expand care management	0.0006 (0.0008)	-0.0005 (0.0008)
Outreach after ED visits and hospitalizations	Expand care management	***0.13**** ***(0.06)***	***0.11**** ***(0.06)***
24/7 patient access to care team	Improve access	***-0.12**** ***(0.06)***	0.00 (0.06)
Same- or next-day access to appointments	Improve access	***0.16**** ***(0.09)***	0.01 (0.07)
Extended hours of operation including weekend	Improve access	-0.03 (0.03)	-0.04 (0.03)

Sources: Mathematica's analysis of practice-reported care delivery data submitted to CMS, CPC+ Practice Survey data, and PBIP performance data from 2017 through 2019 and Medicare Enrollment Database from January 2014 through December 2019.

Note: This table shows the results of 10 regressions, one for each practice capability by track. In each regression, we controlled for baseline practice characteristics, baseline beneficiary characteristics, and current year beneficiary characteristics. All practices were weighted equally, irrespective of their size. Each PBIP score ranges from 0 to 1. Higher PBIP quality scores and PBIP utilization scores indicate better performance (that is, higher quality and lower utilization). The analysis included only non-SSP practices because SSP practices were not eligible to receive PBIPs. ***Yellow shading with bold, italicized text*** signifies that the underlying coefficient was statistically significant at the 10 percent level or better using a two-sided test.

*** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$.

ED = emergency department; PBIP = Performance-based Incentive Payment; QI = quality improvement; SE = standard error; SSP = Medicare Shared Savings Program.

5.P.4. Discussion

The CPC+ model provides a unique opportunity to study performance-based incentive payments for primary care practices and their associations with practice outcomes and actions. Using a rich dataset that we assembled for the CPC+ evaluation, we examined associations between practice capabilities and PBIP scores. We found positive associations between contemporaneous levels of many practice capabilities and PBIP quality scores for both Track 1 and Track 2 practices and positive associations between changes in many practice capabilities and PBIP quality scores for Track 2 practices. We are not aware of any reasons why changes in practice capabilities would translate to higher PBIP quality scores for Track 2 practices but not for Track 1 practices. We found more limited associations between contemporaneous levels of and changes in practice capabilities and PBIP utilization scores, especially for Track 2 practices. These findings are consistent with the takeaway from deep-dive interviews with practices that practices felt they had more control over PBIP quality scores than PBIP utilization scores. Practices also reported implementing more efforts to improve their performance on PBIP quality scores than PBIP utilization scores (Peikes et al. 2021b).

Our analysis of the associations between practices' contemporaneous levels of and changes in practice capabilities and their PBIP scores has several limitations. First, our measures of practice capabilities were self-reported by practices and serve as proxies for CMMI's intended practice activities. Practices that reported a given degree of implementing a practice capability could have implemented the practice capability in substantially different ways (for example, how they calculated patient risk scores and used them to stratify patients by risk level). Second, practices with high levels of practice capabilities are likely different in other ways than practices with low levels of practice capabilities that could help explain differences in PBIP scores. Third, we can only capture impacts in the short run and it might take time for

improvements in practice capabilities to translate to higher PBIP scores. Finally, participation in CPC+ is voluntary and practices in CPC+ are a self-selected group, so the effectiveness of the CPC+ PBIPs and their relationships with practice capabilities might not be generalizable if the model were to be scaled to additional primary care practices in the country.

Nonetheless, the positive associations between practice capabilities and PBIPs found in this analysis are reassuring in that they show that the performance-based payments CPC+ practices received were correlated with the practices' actions (capabilities) in the expected direction. Although the analysis is not causal, the strength of the associations could provide insights for designing and calibrating performance-based payments in future models to incentivize work on practice capabilities and activities that most closely align with policy objectives.

References

American Medical Association (AMA). "Issue Brief: Nation's Drug-Related Overdose and Death Epidemic Continues to Worsen." 2022a. https://www.ama-assn.org/system/files/issue-brief-increases-in-opioid-related-overdose.pdf. Accessed February 21, 2023.

American Medical Association (AMA). "Trends in Health Care Spending." 2022b. https://www.ama-assn.org/about/research/trends-health-care-spending.

Anglin, G., D. Peikes, D. Petersen, A. O'Malley, K. Geonnotti, A. Ghosh, P. Singh, et al. "Independent Evaluation of Comprehensive Primary Care Plus (CPC+): Second Annual Report." Princeton, NJ: Mathematica, 2020.

Attanasio, O., C. Meghir, and M. Szekely. "Using Randomised Experiments and Structural Models for 'Scaling Up': Evidence from the PROGRESA Evaluation." IFS Working Paper EWP03/05, May 2003, p. 43.

Audet, A.-M., and M.A. Zezza. "How Behavioral Economics Can Advance the Design of Effective Clinician Incentive Programs." Health Affairs Blog, September 2015.

Bailit M.H., M.W. Friedberg, and M.L. Houy. "Standardizing the Measurement of Commercial Health Plan Primary Care Spending." New York: Milbank Memorial Fund; 2017.

Baker, L.C., D.P. Kessler, and G.K. Vaska. "The Relationship Between Provider Age and Opioid Prescribing Behavior." *American Journal of Managed Care,* vol. 28, no. 5, May 2022, pp. 223–228.

Barker, P.M., A. Reid, and M. Schall. "A Framework for Scaling Up Health Interventions: Lessons from Large-Scale Improvement Initiatives in Africa." *Implementation Science,* vol. 11, no. 12, 2016.

Billings, J., N. Parikh, and T. Mijanovich. "Emergency Department Use in New York City: A Substitute for Primary Care?" Issue Brief (Commonwealth Fund), vol. 433, 2000, pp. 1–5.

Bohl, A., and M. Roozeboom-Baker. "A COVID-19 Primer: Analyzing Health Care Claims, Administrative Data, and Public Use Files." Princeton, NJ: Mathematica, July 2020. https://www.mathematica.org/our-publications-and-findings/publications/a-COVID-19-primer-analyzing-health-care-claims-administrative-data-and-public-use-files.

Breslau, N., and K.G. Reeb. "Continuity of Care in a University-Based Practice." *Journal of Medical Education and Curricular Development,* vol. 50, no. 10, 1975, pp. 965–969.

Brown, R.S., D. Peikes, G. Peterson, J. Schore, and C.M. Razafindrakoto. "Six Features of Medicare Coordinated Care Demonstration Programs That Cut Hospital Admissions of High-Risk Patients." *Health Affairs*, vol. 31, no. 6, 2012, pp. 1156–1166.

Buchmueller, T.C, and C. Carey. "The Effect of Prescription Drug Monitoring Programs on Opioid Utilization in Medicare." American Economic Journal: Economic Policy, vol. 10, no.1, 2018, pp. 77–112.

Centers for Disease Control and Prevention (CDC). "U.S. Opioid Dispensing Rate Maps." 2019. https://www.cdc.gov/drugoverdose/rxrate-maps/index.html. Accessed September 27, 2021.

Centers for Disease Control and Prevention (CDC). "Prescribing Practices: Changes in Opioid Prescribing Practices." 2020. https://www.cdc.gov/drugoverdose/deaths/prescription/practices.html. Accessed September 27, 2021.

Centers for Disease Control and Prevention (CDC). "Analyzing Opioid Prescription Data and Oral Morphine Milligram Equivalents (MME)." June 2021. https://www.cdc.gov/opioids/data-resources/index.html. Accessed July 30, 2021.

Centers for Medicare & Medicaid Services (CMS). "FY 2022 Hospital Inpatient PPS Final Rule." n.d. https://www.cms.gov/medicare/acute-inpatient-pps/fy-2022-ipps-final-rule-home-page.

Centers for Medicare & Medicaid Services. (CMS). "Public Comment Summary Report. Potential Opioid Overuse." 2018. https://www.cms.gov/Medicare/Quality-Initiatives-Patient-Assessment-Instruments/MMS/Downloads/Electronic-Clinical-Quality-Measure-eCQM-Development-and-Maintenance-for-Eligible-Professionals-EP-eCQM-%E2%80%93-Potential-Opioid-Overuse-Summary-Report.pdf. Accessed October 14, 2021.

Centers for Medicare & Medicaid Services (CMS). "Behavioral Health Integration Services." May 2019a. https://www.cms.gov/Outreach-and-Education/Medicare-Learning-Network-MLN/MLNProducts/Downloads/BehavioralHealthIntegration.pdf. Accessed October 30, 2019.

Centers for Medicare and Medicaid Services (CMS), Center for Medicare and Medicaid Innovation. "Comprehensive Primary Care Initiative." 2019b. https://innovation.cms.gov/initiatives/comprehensive-primary-care-initiative. Accessed December 10, 2019.

Centers for Medicare & Medicaid Services. (CMS). "CMS Roadmap: Strategy to Fight the Opioid Crisis." 2020. https://www.cms.gov/About-CMS/Agency-Information/Emergency/Downloads/Opioid-epidemic-roadmap.pdf. Accessed March 3, 2022. (5G)

Chee, T.T, A. Ryan, J.H. Wasfy, and W.B. Borden. "Current State of Value-Base Purchasing Programs." *Circulation,* vol. 133, no. 22, 2016, pp. 2197–2205.

Chen, J.H., K. Humphreys, N.H. Shah, and A. Lembke. "Distribution of Opioids by Different Types of Medicare Prescribers." *JAMA Internal Medicine,* vol. 176, no. 2, 2016, pp. 259–261. https://doi.org/10.1001/jamainternmed.2015.6662.

Chhabra, K.R., U. Nuliyalu, J.B. Dimick, and H. Nathan. "Who Will Be the Costliest Patients? Using Recent Claims to Predict Expensive Surgical Episodes." *Medical Care,* vol. 57, no. 11, November 2019, pp. 869–874. https://doi.org/10.1097/mlr.0000000000001204.

Chou, R., D. Hartung, J. Turner, I. Blazina, B. Chan, X. Levander, M. McDonagh, S. Selph, R. Fu, and M. Pappas. "Opioid Treatments for Chronic Pain." Comparative Effectiveness Review No. 229. Agency for Healthcare Research and Quality, April 2020. https://doi.org/10.23970/AHRQEPCCER229.

Cole, S., and E. Stuart. "Generalizing Evidence from Randomized Clinical Trials to Target Populations: The ACTG 320 Trial." *American Journal of Epidemiology,* vol. 172, no. 1, July 2010, pp. 107–115.

Congressional Research Service (CRS). "Consumption of Prescription Opioids for Pain: A Comparison of Opioid Use in the United States and Other Countries." June 2, 2021. https://crsreports.congress.gov/product/pdf/R/R46805. Accessed February 23, 2022.

Correa, J., J. Tian, and E. Bareinboim. "Generalized Adjustment Under Confounding and Selection Biases." Presentation at the AAAI Conference on Artificial Intelligence, 2018, pp. 6335–6342.

Crabtree B.F., P.A. Nutting, W.L. Miller, R.R. McDaniel, K.C. Stange, C.R. Jaen, and E. Stewart. "Primary Care Practice Transformation Is Hard Work: Insights from a 15-Year Developmental Program of Research." *Medical Care,* vol. 49 (suppl.), 2011, pp. S28–35.

Cramer, J.D., V. Gunaseelan, H.M. Hu, M.C. Bicket, J.F. Waljee, and M.J. Brenner. "Association of State Opioid Prescription Duration Limits with Changes in Opioid Prescribing for Medicare Beneficiaries." *JAMA Internal Medicine,* vol. 181, no. 12, 2021, pp. 1656–1657. https://doi.org/10.1001/jamainternmed.2021.4281.

Dahabreh, I.J., S.E. Robertson, J.A. Steingrimsson, E.A. Stuart, and M.A. Hernan. "Extending Inferences from a Randomized Trial to a New Target Population." *arXiv,* 1805.00550, 2018.

Dale, S.B., A. Ghosh, D.N. Peikes, T.J. Day, F.B. Yoon, E.F. Taylor, K. Swankoski, et al. "Two-Year Costs and Quality in the Comprehensive Primary Care Initiative." *New England Journal of Medicine,* vol. 374, no. 24, 2016, pp. 2345–2356.

Daw, J.R., and L.A. Hatfield. "Matching and Regression to the Mean in Difference-in-Differences Analysis." *Health Services Research,* vol. 53, no. 6, 2018, pp. 4138–4156. doi:10.1111/1475-6773.12993.

Degtiar, I., and S. Rose. "A Review of Generalizability and Transportability." *Annual Review of Statistics and Its Application,* vol. 10, no. 1, 2023. https://doi.org/10.1146/annurev-statistics-042522-103837.

Delgado, M., F.S. Shofer, M.S. Patel, S. Halpern, C. Edwards, Z.F. Meisel, and J. Perrone. "Association Between Electronic Medical Record Implementation of Default Opioid Prescription Quantities and Prescribing Behavior in Two Emergency Departments." *Journal of General Internal Medicine,* vol. 33, no. 4, 2018, pp. 409–411. https://doi.org/10.1007/s11606-017-4286-5.

Dowell, D., T.M. Haegerich, and R. Chou. "CDC Guideline for Prescribing Opioids for Chronic Pain—United States, 2016." *JAMA*, vol. 315, no. 15, 2016, pp.1624–1645. https://doi.org/10.1001/jama.2016.1464.

Dowell, D., K. Ragan, C.M. Jones, G. Baldwin, and R. Chou. "Prescribing Opioids for Pain—The New CDC Clinical Practice Guideline." *New England Journal of Medicine,* vol. 387, 2022, pp. 2011–2013. https://doi.org/10.1056/NEJMp2211040.

Duan, L., M.-S. Lee, J.L. Adams, A.L. Sharp, and J.S. Doctor. "Opioid and Naloxone Prescribing Following Insertion of Prompts in the Electronic Health Record to Encourage Compliance with California State Opioid Law." *JAMA Network Open,* vol. 5, no. 5, 2022, e229723. https://doi.org/10.1001/jamanetworkopen.2022.9723.

eCQI Resource Center. "Potential Opioid Overuse." n.d. https://ecqi.healthit.gov/sites/default/files/ecqm/measures/CMS460v2.html. Accessed July 30, 2021.

Els, C., T.D. Jackson, D. Kunyk, V.G. Lappi, B. Sonnenberg, R. Hagtvedt, S. Sharma, et al. "Adverse Events Associated with Medium- and Long-Term Use of Opioids for Chronic Non-Cancer Pain: An Overview of Cochrane Reviews." *Cochrane Database of Systematic Reviews,* vol. 10, 2017, CD012509. https://doi.org/10.1002/14651858.CD012509.pub2. Accessed October 14, 2021.

Finkel, J., and D. Marriott. "What Makes Michigan's High-Performing Primary Care Practices Work Well?" Milbank Memorial Fund (blog), January 12, 2021. https://www.milbank.org/2021/01/what-makes-michigans-high-performing-primary-care-practices-work-well/. Accessed January 18, 2022.

Fleming, C., E. Shin, R. Powell, et al. "Updating a Claims-Based Measure of Low-Value Services Applicable to Medicare Fee-for-Service Beneficiaries." *Journal of General Internal Medicine.* Published online June 6, 2022. https://doi.org/10.1007/s11606-022-07654-7.

Flores, C.A., and O.A. Mitnik. "Comparing Treatments Across Labor Markets: An Assessment of Nonexperimental Multiple-Treatment Strategies." *The Review of Economics and Statistics, vol.* 95, no. 5, 2013, pp. 1691–1707. https://doi.org/10.1162/REST_a_00373.

Freese, K.E., A. Vega, J.J. Lawrence, and P.I. Documet. "Social Vulnerability Is Associated with Risk of COVID-19 Related Mortality in US Counties with Confirmed Cases." *Journal of Health Care for the Poor and Underserved,* vol. 32, no. 1, 2021, pp. 245–257. doi: 10.1353/hpu.2021.0022. https://muse.jhu.edu/article/783108

Fu, N., P. Singh, S. Dale, S. Orzol, D. Peikes, A. Ghosh, R. Brown, and T.J. Day. "Long-Term Effects of the Comprehensive Primary Care Model on Health Care Spending and Utilization." *Journal of General Internal Medicine,* vol. 37, no. 7, 2022, pp. 1713–1721.

Gechter, M. "Generalizing the Results from Social Experiments: Theory and Evidence from Mexico and India." Department of Economics, Pennsylvania State University, unpublished manuscript, 2015.

Gelman, A., and T.C. Little. "Poststratification into Many Categories Using Hierarchical Logistic Regression." *Survey Methodology,* vol. 23, 1997, pp. 127–135.

Ghosh, A., S. Orzol, S. Dale, J. Laird, N. Fu, P. Singh, and M.-Y. Kim, et al. "Independent Evaluation of Comprehensive Primary Care Plus (CPC+): Second Annual Report, Appendices to the Supplemental Volume." Mathematica, July 2020.

Green, D., and H. Kern. "Modeling Heterogeneous Treatment Effects in Survey Experiments with Bayesian Additive Regression Trees." *Public Opinion Quarterly,* vol. 76, no. 3, 2012, pp. 491–511.

Haas, A., M.N. Elliott, J.W. Dembosky, J.L. Adams, S.M. Wilson-Frederick, J.S. Mallett, S. Gaillot, et al. 2019. "Imputation of Race/Ethnicity to Enable Measurement of HEDIS Performance by Race/Ethnicity." *Health Services Research,* vol. 54, no. 1, pp. 13–23. https://doi.org/10.1111/1475-6773.13099.

Hadland, S.E., A. Rivera-Aguirre, B.D.L. Marshall, and M. Cerdá. "Association of Pharmaceutical Industry Marketing of Opioid Products With Mortality From Opioid-Related Overdoses." *JAMA Network Open,* vol. 2, no. 1, 2019, e186007. https://doi.org/10.1001/jamanetworkopen.2018.6007.

Hahn, P., J. Murray, and C. Carvalho. "Bayesian Regression Tree Models for Causal Inference: Regularization, Confounding, and Heterogeneous Effects." *arXiv,* vol. 1706.09523v3, 2019, pp. 1–33.

Hartman, M., A.B. Martin, N. Espinosa, A. Catlin, and National Health Expenditure Accounts Team. "National Health Care Spending in 2016: Spending and Enrollment Growth Slow After Initial Coverage Expansions." *Health Affairs,* vol. 37, no. 1, 2018, pp. 150–160.

Hill, J. "Bayesian Nonparametric Modeling for Causal Inference." *Journal of Computational and Graphical Statistics,* vol. 20, no. 1, 2011, pp. 217–240.

Hollander, M.A.G., J.M. Donohue, B.D. Stein, E.E. Krans, and M.P. Jarlenski. "Association Between Opioid Prescribing in Medicare and Pharmaceutical Company Gifts by Physician Specialty." *Journal of General Internal Medicine,* vol. 35, no. 8, 2020, pp. 2451–2458. https://doi.org/10.1007/s11606-019-05470-0.

Humphreys, K., and J. Marsden. "Synthesizing Advice to the United States from Other Nations That Have Experienced Drug Epidemics." *Addiction,* vol 117, no.5, 2022, pp. 1202–1203.

Islam, N., S.J. Sharp, G. Chowell, S. Shabnam, I. Kawachi, B. Lacey, J.M. Massaro, et al. "Physical Distancing Interventions and Incidence of Coronavirus Disease 2019: Natural Experiment in 149 Countries." *BMJ,* vol. 370, 2020, m2743. doi: https://doi.org/10.1136/bmj.m2743.

Islam, S., A. Nayak, Y. Hu, A. Mehta, K. Dieppa, Z. Almuwaqqat, Y. Ko, et al. "Temporal Trends in the Association of Social Vulnerability and Race/Ethnicity with County-Level COVID-19 Incidence and Outcomes in the USA: An Ecological Analysis." *BMJ Open,* vol. 11, 2021, e048086. doi:10.1136/bmjopen-2020-048086.

Johnston, K. J., L. Allen, T. A. Melanson, and S. R. Pitts. "A Patch to the NYU Emergency Department Visit Algorithm." *Health Services Research*, vol. 52, no. 4, 2017, pp. 1264–1276.

Karaye, I.M., and J.A. Horney. "The Impact of Social Vulnerability on COVID-19 in the U.S.: An Analysis of Spatially Varying Relationships." *American Journal of Preventive Medicine,* vol. 59, no. 3, 2020. doi: 10.1016/j.amepre.2020.06.006.

Katcher, B.A., and C.J. Ruhm. "Federal Opioid Grant Funding Favors Least Populous States, not States with Greatest Need: Study Examines how Closely Federal Grant Dollars Awarded in Fiscal Years 2017–2018 Aligned with the Severity of State Opioid Problems." *Health Affairs,* vol. 40, no. 5, 2021, pp. 829–836.

Kempski A., and A. Greiner. "Primary Care Spending: High Stakes, Low Investment." Washington, DC: Primary Care Collaborative, 2020.

Kern, H., E. Stuart, J. Hill, and D. Green. "Assessing Methods for Generalizing Experimental Impact." *Journal of Research on Educational Effectiveness,* vol. 9, no. 1, 2016, pp. 103–127.

Kern, L.M., J. Seirup, L.P. Casalino, and M.M. Safford. "Healthcare Fragmentation and the Frequency of Radiology and Other Diagnostic Tests." *Journal of General Internal Medicine,* vol. 32, no. 2, 2017, pp. 175–181.

Kerns, R.D. "Social and Behavioral Sciences: Response to the Opioid and Pain Crises in the United States." *American Journal of Public Health,* vol. 112, no. S1, 2022, pp. S6–S8.

Kerr, E.E., E.A. McGlynn, K.A. VanVoorst, et al. "Measuring Antidepressant Prescribing Practice in a Health Care System Using Administrative Data: Implications for Quality Measurement and Improvement." *Joint Commission Journal on Quality Improvement,* vol. 24, 2000, pp. 203–216.

Keshwani, S., M. Maguire, A. Goodin, W. Lo-Ciganic, D.L. Wilson, and J.M. Hincapie-Castillo. "Buprenorphine Use Trends Following Removal of Prior Authorization Policies for the Treatment of Opioid Use Disorder in 2 State Medicaid Programs." *JAMA Health Forum,* vol. 3, no. 6, 2022, e221757. https://doi.org/10.1001/jamahealthforum.2022.1757.

Khuller, D., D.A. Chokshi, R. Kocher, A. Reddy, K. Basu, P.H. Conway, and R. Rajkumar. "Behavioral Economics and Physician Compensation—Promise and Challenges." *New England Journal of Medicine,* vol. 372, 2015, pp. 2281–2283.

Kranker, K., L. Blue, and L.V. Forrow. "Improving Effect Estimates by Limiting the Variability in Inverse Propensity Score Weights." *American Statistician,* vol. 75, no. 3, 2021, pp. 276–287. https://doi.org/10.1080/00031305.2020.1737229.

Laird, J., A. Markovitz, J. Zurovac, P. Singh, D. Petersen, P. Shanmugam, N. Morrison, et al. "Independent Evaluation of Comprehensive Primary Care Plus (CPC+): Appendices to the Fourth Annual Report." Princeton, NJ: Mathematica, May 2022. https://innovation.cms.gov/data-and-reports/2022/cpc-plus-fourth-annual-eval-report-app.

Lehmann, S.W., and M. Fingerhood. "Substance-Use Disorders in Later Life." *New England Journal of Medicine,* vol. 379, no. 24, 2018, pp. 2351–2360. https://doi.org/10.1056/NEJMra1805981.

Lei, L., E. Rich, and A. Ghosh. "Pandemic-Related Physician and Practice Site Service Interruptions in CPC+ and Comparison Practice Sites." (Memo.) September 9, 2022.

Liang, L., B. Moore, and A. Soni. "National Inpatient Hospital Costs: The Most Expensive Conditions by Payer, 2017." In *Healthcare Cost and Utilization Project (HCUP) Statistical Briefs.* Agency for Healthcare Research and Quality, July 2020.

Liebschutz, J.M., Z. Xuan, C.W. Shanahan, M. LaRochelle, J. Keosaian, D. Beers, G. Guara, et al. "Improving Adherence to Long-Term Opioid Therapy Guidelines to Reduce Opioid Misuse in Primary Care: A Cluster-Randomized Clinical Trial." *JAMA Internal Medicine,* vol. 177, no. 9, 2017, pp.1265–1272. https://jamanetwork.com/journals/jamainternalmedicine/fullarticle/2643777.

Lo-Ciganic, W.H., J.L. Huang, H.H. Zhang, J.C. Weiss, Y. Wu, C.K. Kwoh, J.M. Donohue, et al. "Evaluation of Machine-learning Algorithms for Predicting Opioid Overdose Risk among Medicare Beneficiaries with Opioid Prescriptions." *JAMA Network Open,* vol. 2, no. 3, 2019, pp. e190968–e190968.

Maratt, J.K., E.A. Kerr, M.L. Klamerus, et al. "Measures Used to Assess the Impact of Interventions to Reduce Low-Value Care: A Systematic Review." *Journal of General Internal Medicine,* vol. 34, no. 9, 2019, pp. 1857–1864.

March, J.P., J.Y. Lim, K.L. Manzione, M. Buncke, and M.L. Shindo. "Association of a Multimodal Intervention with Decreased Opioid Prescribing After Neck Dissection for Malignant Thyroid Disease with Short Hospital Stay." *JAMA Otolaryngology Head & Neck Surgery,* vol. 148, no. 6, 2022, pp. 561–567. https://doi.org/10.1001/jamaoto.2022.0952.

McNellis, R.J., J.L. Genevro, and D.S. Meyers. "Lessons Learned from the Study of Primary Care Transformation." *Annals of Family Medicine,* vol. 11, suppl. 1, 2013, pp. S1–S5.

Meisenberg, B.R., J. Grover, C. Campbell, and D. Korpon. "Assessment of Opioid Prescribing Practices Before and After Implementation of a Health System Intervention to Reduce Opioid Overprescribing." *JAMA Network Open,* vol. 1, no. 5, 2018, e182908. https://doi.org/10.1001/jamanetworkopen.2018.2908.

National Committee for Quality Assurance (NCQA). "HEDIS Volume 2: Technical Specifications." 2015. http://www.ncqa.org/HEDISQualityMeasurement/HEDISMeasures/HEDIS2015.aspx.

National Committee for Quality Assurance (NCQA). "HEDIS Volume 2: Technical Specifications." 2018. http://www.ncqa.org/hedis-quality-measurement/hedis-measures/hedis-2018.

National Committee for Quality Assurance (NCQA). "HEDIS MY 2020 Medication List Directory." 2020. https://store.ncqa.org/hedis-my-2020-medication-list-directory.html. Accessed July 30, 2021.

National Committee for Quality Assurance (NCQA). "HEDIS MY 2021 Medication List Directory." 2021. https://store.ncqa.org/hedis-my-2021-medication-list-directory.html. Accessed July 30, 2021.

National Committee for Quality Assurance (NCQA). "HEDIS MY 2022 Medication List Directory." 2022. https://store.ncqa.org/hedis-my-2022-medication-list-directory.html. Accessed July 30, 2022.

Noyes, K., L. Hangsheng, J.M. Lyness, and B. Friedman. "Medicare Beneficiaries with Depression: Comparing Diagnoses in Claims Data with the Results of Screening." *Psychiatric Services,* vol. 62, no. 10, 2011, pp. 1159–1165.

Nutting, P.A., W.L. Miller, B.F. Crabtree, C.R. Jaen, E.E. Stewart, and K.C. Stange. "Initial Lessons from the First National Demonstration Project on Practice Transformation to a Patient-Centered Medical Home." *Annals of Family Medicine,* vol. 7, no. 3, 2009, pp. 254–260.

Nyweide, D.J., D.L. Anthony, J.P. Bynum, R.L. Strawderman, W.B. Weeks, L.P. Casalino, and E.S. Fisher. "Continuity of Care and the Risk of Preventable Hospitalization in Older Adults." *JAMA Internal Medicine,* vol. 173, 2013, pp. 1879–1885.

O'Malley, A.S., E.C. Rich, L. Shang, T. Rose, A. Ghosh, D. Poznyak, and D. Peikes. "New Approaches to Measuring the Comprehensiveness of Primary Care Physicians." *Health Services Research,* vol. 54, no. 2, 2019, pp. 356–366.

O'Malley, A.S., E.C. Rich, L. Shang, T. Rose, A. Ghosh, D. Poznyak, D. Peikes, and M. Niedzwiecki. "Practice-site-level Measures of Primary Care Comprehensiveness and Their Associations with Patient Outcomes." *Health Services Research,* vol. 56, no. 3, 2021, pp. 371–377.

O'Malley, A., P. Singh, N. Fu, N. Duda, N. McCall, K. Geonnotti, D. Petersen, et al. "Independent Evaluation of Comprehensive Primary Care Plus (CPC+): Final Annual Report." Princeton, NJ: Mathematica, 2023.

Orzol, S., K. Geonnotti, R. Keith, I. Degtiar, D. Peterson, A. O'Malley, N. Duda, et al. "Independent Evaluation of Comprehensive Primary Care Plus (CPC+): Design Report." Princeton, NJ: Mathematica. Unpublished, last modified May 2022.

Orzol, S., A. Ghosh, N. Fu, J. Laird, S. Dale, L. Timmins, K. Swankoski, et al. "Independent Evaluation of Comprehensive Primary Care Plus (CPC+): Appendices to the Third Annual Report." Princeton, NJ: Mathematica, 2021.

Oster, A.M., G.J. Kang, A.E. Cha, V. Beresovsky, C.E. Rose, G. Rainisch, and L. Porter. "Trends in Number and Distribution of COVID-19 Hotspot Counties—United States, March 8–July 15, 2020." *Morbidity and Mortality Weekly Report,* vol. 69, no. 33, 2020, pp. 1127–1132. doi: http://dx.doi.org/10.15585/mmwr.mm6933e2.

Parchman, M.L., R.B. Penfold, B. Ike, D. Tauben, M. Von Korff, M. Stephens, K.A. Stephens, et al. "Team-Based Clinic Redesign of Opioid Medication Management in Primary Care: Effect on Opioid Prescribing." *Annals of Family Medicine,* vol. 17, no. 4, 2019, pp. 319–325.

Pardo, B. "Do More Robust Prescription Drug Monitoring Programs Reduce Prescription Opioid Overdose?" *Addiction,* vol. 11, no. 10, 2017, 1773–1783. https://onlinelibrary.wiley.com/doi/epdf/10.1111/add.13741.

Peikes, D., S. Orzol, K. Geonnotti, D. Petersen, A. O'Malley, N. Duda, S. Dale, et al. "Independent Evaluation of Comprehensive Primary Care Plus (CPC+) Draft Design Report, Appendices: Appendix 5.B. Design for CPC+ Causal Path Analysis for ED Visits and Inpatient Visits." Princeton, NJ: Mathematica Policy Research, 2021a.

Peikes, D., K. Swankoski, A. O'Malley L. Timmins, D. Petersen, K. Geonnotti, H. Tu, et al. "Independent Evaluation of Comprehensive Primary Care Plus (CPC+): Third Annual Report." Mathematica, January 2021b.

Peikes, D., E. Fries Taylor, A.S. O'Malley, and E.C. Rich. "The Changing Landscape of Primary Care: Effects of the ACA and Other Efforts Over the Past Decade." *Health Affairs,* vol. 9, no. 3, 2020, pp. 421–428.

Peterson, G., S. Dale, and D. Peikes. "Primary and Secondary Tests for the CPC+ Impact Evaluation." Memorandum to CMS. Princeton, NJ: Mathematica Policy Research, May 23, 2018.

Pharmacy Quality Alliance (PQA). "PQA Quality Measures." 2021. Available at https://www.pqaalliance.org/pqa-measures. Accessed May 15, 2022.

Pollack, C.E., P.S. Hussey, R.S. Rudin, D.S. Fox, J. Lai, and E.C. Schneider. "Measuring Care Continuity: A Comparison of Claims-Based Methods." *Medical Care,* vol. 54, no. 5, 2016, pp. e30–e34.

Polyakova, M., V. Udalova, G., Kocks K. Genadek, K. Finlay, and A. Finkelstein. "Racial Disparities in Excess All-Cause Mortality During the Early COVID-19 Pandemic Varied Substantially Across States." *Health Affairs,* vol. 40, no. 2, 2020, pp. 307–316. https://doi.org/10.1377/hlthaff.2020.02142.

Pope, G.C., J. Kautter, R.P. Ellis, A.S. Ash, J.Z. Ayanian, L.I. Iezzoni, M.J. Ingber, et al. "Risk Adjustment of Medicare Capitation Payments Using the CMS-HCC Model." *Health Care Financing Review,* vol. 25, no. 4, 2004, pp. 119–141.

Pope, G.C., J. Kautter, M.J. Ingber, S. Freeman, R. Sekar, and C. Newhart. "Evaluation of the CMS-HCC Risk Adjustment Model. Final Report." Research Triangle Park, NC: RTI International, March 2011.

Powell, B.J., T.J. Waltz, M.J. Chinman, L. Damschroder, J.L. Smith, M.M. Matthieu, E.K. Proctor, and J.E. Kirchner. "A Refined Compilation of Implementation Strategies: Results from the Expert Recommendations for Implementing Change (ERIC) Project." *Implementation Science,* vol. 10, no. 21, 2015.

Prescription Drug Monitoring Program Training and Technical Assistance Center (PDMP TTAC). "Maps and Tables." 2021. https://www.pdmpassist.org/Policies/Maps. Accessed September 28, 2021.

Reid R., C. Damberg, and M.W. Friedberg. "Primary Care Spending in the Fee-for-Service Medicare Population." *JAMA Internal Medicine*, vol. 179, no. 7, 2019, pp. 977–980.

Rich, E., D. Lipson, J. Libersky, and M. Parchman. "Coordinating Care for Adults with Complex Care Needs in the Patient-Centered Medical Home: Challenges and Solutions." Agency for Healthcare Research and Quality (AHRQ) Publication No. 12-0010-EF. Rockville, MD: AHRQ, January 2012.

Rich E.C., A.S. O'Malley, C. Burkhart, et al. "Report on Measuring Range of Services of Physicians and of Practices in CPC+: Implications for CPC+ and PCF Evaluations." Memo developed for CMS. January 2021.

Rudolph, K., and M. van der Laan. "Robust Estimation of Encouragement Design*." Journal of the Royal Statistical Society Series B Statistical Methodology,* vol. 79, no. 5, 2017, pp. 1509–1525.

Schepis, T.S., D.L. Klare, J.A. Ford, and S.E. McCabe. "Prescription Drug Misuse: Taking a Lifespan Perspective." *Substance Abuse: Research and Treatment,* vol. 14, 2020. https://doi.org/10.1177/1178221820909352.

Schommer, J., S. Allen, N. Scholz, M. Reams, and D. Bohn. "Evaluation of Quality Improvement Methods for Altering Opioid Prescribing Behavior in Hand Surgery." *Journal of Bone and Joint Surgery,* American Volume, vol. 102, no. 9, 2020, pp. 804–810. https://doi.org/10.2106/JBJS.19.01052.

Schwartz, A.L., M.E. Chernew, B.E. Landon, and J.M. McWilliams. "Changes in Low-Value Services in Year 1 of the Medicare Pioneer Accountable Care Organization Program." *JAMA Internal Medicine,* vol. 175, no. 11, 2015, pp. 1815–1825.

Schwartz, A.L., B.E. Landon, A.G. Elshaug, M.E. Chernew, and J.M. McWilliams. "Measuring Low-Value Care in Medicare." *JAMA Internal Medicine*, vol. 174, no. 7, 2014, pp. 1067–1076.

Seal, K.H., T. Rife, Y. Li, C. Gibson, and J. Tighe. "Opioid Reduction and Risk Mitigation in VA Primary Care: Outcomes from the Integrated Pain Team Initiative." *Journal of General Internal Medicine,* vol. 35, 2020, pp.1238–1244. https://doi.org/10.1007/s11606-019-05572-9.

Singh, P., S. Orzol, D. Peikes, E. Oh, and S. Dale. "Characteristics of Practices that Participate in a Primary Care Transformation Model: Evidence from the Comprehensive Primary Care Plus Initiative." *Annals of Family Medicine,* vol. 18, 2020, pp. 309–317.

Solberg, L., K. Engebreston, J.M. Sperl-Hillen, et al. "Are Claims Data Accurate Enough to Identify Patients for Performance Measures or Quality Improvement? The Case of Diabetes, Heart Disease and Depression." *American Journal of Medical Quality,* vol. 21, no. 4, 2006, pp. 238–245.

Song, Z. "Mortality Quadrupled Among Opioid-Driven Hospitalizations, Notably Within Lower-Income and Disabled White Populations." *Health Affairs,* vol. 36, no. 12, 2017, pp. 2054–2061.

Swankoski, K., A. O'Malley, H. Tu, D. Petersen, P. Singh, K. Geonnotti, R. Keith, et al. "Independent Evaluation of Comprehensive Primary Care Plus (CPC+): Fourth Annual Report." Mathematica, 2022.

Tamblyn, R., N. Girard, J. Boulet, D. Dauphinee, and B. Habib. "Association of Clinical Competence, Specialty and Physician Country of Origin with Opioid Prescribing for Chronic Pain: A Cohort Study." *BMJ Quality & Safety,* vol. 31, 2022, pp. 340–352.

Templeman, M.L., L. Field, W.A. Jones, M. Stanley, A. Bohl, and S. Ode. "Restructured BETOS Classification System (RBCS) Final Report." Erie, PA: Provider Resources, Inc, September 2020.

Timmins, L., D. Peikes, and N. McCall. "Pathways to Reduced Emergency Department and Urgent Care Center Use: Lessons from the Comprehensive Primary Care Initiative." *Health Services Research,* vol. 55, no. 6, 2020, pp. 1003–1012.

Townsend, L., J.T. Walkup, S. Crystal, and M. Olfson. "A Systematic Review of Validated Methods for Identifying Depression Using Administrative Data." *Pharmacoepidemiology and Drug Safety,* vol. 21, suppl. 1, 2012, pp. 163–173. https://doi.org/10.1002/pds.2310.

Urgent Care Association of America. Urgent Care Fast Facts. n.d. https://urgentcareassociation.org/.

Urgent Care Association of America. *2019 Benchmarking Report*. 2021.

Von Korff, M., A. Kolodny, R.A. Deyo, and R. Chou. "Long-Term Opioid Therapy Reconsidered." *Annals of Internal Medicine,* vol. 155, 2011, 325–328. https://www.acpjournals.org/doi/pdf/10.7326/0003-4819-155-5-201109060-00011.

Wasserstein, R.L., and N.A. Lazar. "The ASA Statement on P-Values: Context, Process, and Purpose." *American Statistician,* vol. 70, no. 2, 2016, pp. 129–133. https://doi.org/10.1080/00031305.2016.1154108.

Wen, H., B.R. Schackman, B. Aden, and Y. Bao. "States with Prescription Drug Monitoring Mandates Saw a Reduction in Opioids Prescribed to Medicaid Enrollees." *Health Affairs,* vol. 36, no. 4, 2017, 733–741. https://www.healthaffairs.org/doi/pdf/10.1377/hlthaff.2016.1141.

World Health Organization. "Nine Steps for Developing a Scaling-Up Strategy." 2010. https://apps.who.int/iris/handle/10665/44432. Accessed February 2, 2023.

Yale New Haven Health Services Corporation/Center for Outcomes Research & Evaluation (YNHHSC/CORE). "2019 Hospital-Wide Readmission Measure Updates and Specifications Report – Version 8.0." Prepared for the Centers for Medicare & Medicaid Services, March 2019.

Yale New Haven Health Services Corporation/Center for Outcomes Research & Evaluation (YNHHSC/CORE). "2021 Hospital-Wide Readmission Measure Updates and Specifications Report – Version 10.0." Prepared for the Centers for Medicare & Medicaid Services, April 2021. (

Yoshikawa, A., G. Ramirez, M.L. Smith, M. Foster, A. K. Nabil, S.N. Jani, and M.G. Ory. "Opioid Use and the Risk of Falls, Fall Injuries and Fractures Among Older Adults: A Systematic Review and Meta-Analysis." *Journals of Gerontology: Series A, Biological Sciences and Medical Sciences,* vol. 75, no. 10, 2020, pp. 1989–1995. https://doi.org/10.1093/gerona/glaa038.

Yousman, L.C., A. Khunte, W. Hsiang, S. Jain, H. Forman, and D. Wiznia. "Urgent Care Center Wait Times Increase for COVID-19 Results in August 2020, with Rapid Testing Availability Limited." *BMC Health Services Research*, vol. 21, no. 1, 2021, pp. 1–5.

Mathematica Inc.

Princeton, NJ • Ann Arbor, MI • Cambridge, MA
Chicago, IL • Oakland, CA • Seattle, WA
Woodlawn, MD • Washington, DC

Mathematica®
Progress Together

mathematica.org website

EDI Global, a Mathematica Company

Operating in Tanzania, Uganda, Kenya, Mozambique, and the United Kingdom