Evaluation of the Independence at Home Demonstration

An Examination of Year 7, the First Year of the COVID-19 Pandemic – Appendices

January 2023

Laura Kimmey, Jason Rotter, Joseph Lovins, Rachel Kogan

Contributing authors: Linda Barterian, Claire Burkhardt, Valerie Cheh, Christine Cheu, Irna May Connor, Tyler Fisher, Margaux Flemming, Lauren Vollmer Forrow, David Jones, Scott Ode, Maya Palakal, Tyler Rose, Beny Wu

Submitted to:

Centers for Medicare & Medicaid Services
Center for Medicare & Medicaid Innovation
7500 Security Blvd.
Baltimore, MD 21244
Project Officer: David Nyweide
Contract Number:
HHSM-500-2014-00034I/HHSM-500-T0018

Submitted by:

Mathematica
1100 First Street, NE, 12th Floor
Washington, DC 20002-4221
Telephone: (202) 484-9220
Facsimile: (202) 863-1763
Project Director: Laura Kimmey

Contents

Exhibits

Appendix A: Technical Appendix

1. Overview

Congress mandated the Independence at Home (IAH) demonstration to test a payment incentive and service delivery model for home-based primary care. Under the IAH demonstration, physicians and nurse practitioners (NPs) direct home-based primary care teams. These teams aim to reduce expenditures and improve the health outcomes of Medicare beneficiaries with multiple chronic conditions and substantial functional limitations. As we discussed in Chapter 1 of the report, the IAH demonstration introduced an incentive to reduce Medicare expenditures (incentive payments) and a service delivery model (home-based primary care led by physicians or NPs). As we described in Chapter 2 of the report, the Mathematica study team estimated a difference-in-differences model to determine whether the demonstration affected Medicare expenditures and hospital use. We also examined whether IAH affected probability of death or entry into institutional long-term care. In this appendix, we present the sample, data, and methods we used for these quantitative analyses as well as analyses of claims data in Chapter 3 of the report to understand how the practices provided care during the COVID-19 pandemic.

The quantitative evaluation design of the demonstration was a difference-in-differences analysis using repeated cross-sections of eligible beneficiaries within demonstration practices (which we also refer to as sites) with a propensity score-matched comparison group that did not receive home-based primary care. We had two years of pre-demonstration data and seven years of post-demonstration data (that is, the first seven years of the demonstration). We observed beneficiaries for the number of months they were eligible for IAH for each demonstration year. To determine the effect of the demonstration on expenditures (and other outcomes) in a given year, such as Year 7, we did the following:

- Estimated the difference in Medicare expenditures per beneficiary per month (PBPM) between the year before the demonstration (the baseline year) and Year 7 for IAH beneficiaries. We restricted claims to those occurring between the date of eligibility for the demonstration in a given year and the end of that year (or date of death).[1] We controlled for beneficiaries' characteristics, such as time since most recent hospital admission; demographic characteristics; activities of daily living (ADLs); and several measures of health status, including the Centers for Medicare & Medicaid Services (CMS) Hierarchical Condition Categories (HCC) risk score. We provide a complete list of control variables later in this appendix.

[1] As we discuss in Chapter 5 of this appendix, we restricted claims the same way in all pre-demonstration years and demonstration years. For example, if a beneficiary became eligible for IAH in month 3 of a given year, we count their expenditures for months 3 to 12. As we discuss in Chapter 4, about two-thirds of beneficiaries become eligible on the first day of each year.

- Estimated the difference in Medicare expenditures during the same period for comparison beneficiaries. As with the IAH group, we restricted claims to those that occurred between the date of eligibility and the end of the year, controlling for beneficiary characteristics.
- Obtained the estimated effect of the demonstration by calculating the difference between the change in expenditures for IAH beneficiaries and the change in expenditures for comparison beneficiaries.

We refer to this model as a difference-in-differences model because it measured the change between two differences (differences for the IAH beneficiaries from the year before the demonstration and the year of interest, such as Year 7, and the difference over the same time period for the comparison group). This method estimated the effect of the demonstration by accounting for two factors. First, it accounted for the difference in expenditures between IAH and comparison beneficiaries before the demonstration. Second, it accounted for changes in expenditures during the demonstration caused by factors unrelated to the demonstration that affected IAH and comparison beneficiaries equally. The difference-in-differences design provides a strong assessment of the demonstration's effect, assuming that the difference in expenditures between IAH and comparison beneficiaries was stable before the demonstration. As we describe later, we tested this assumption.

Interpreting the effect of the demonstration

The COVID-19 pandemic changed how we interpret results of the evaluation, since Year 7 coincided with the first year of the pandemic (2020). In Year 7, the estimated effects reflect both of the following during the first year of the pandemic:

- Any effects of changes in care delivery by IAH practices because of the IAH payment incentive, which was the focus of the evaluation in Years 1 to 6.
- Any changes in the relative effectiveness of home-based primary care for IAH beneficiaries.

The difference-in-differences model, however, is not without limitations; we address our evaluation's limitations in Chapter 4 of the report. Importantly, the COVID-19 pandemic changed how we interpret the estimated effect of IAH; Year 7 coincided with the first year of the pandemic (2020), and the pandemic may have affected the IAH and comparison groups in different ways that we could not measure. For more information, see Chapter 6 of this appendix and Chapters 1 and 4 of the report.

Eighteen practices began the demonstration in 2012. Our total sample for estimating the impact of the demonstration consisted of 14 practices in Years 1 to 5, 12 practices in Year 6, and 10 practices in Year 7, counting the consortium in Richmond as one practice (Exhibit A.1). In all years, our quantitative analyses excluded three practices (Atlanta, Chicago, and Stuart) that withdrew from the demonstration before Year 3 and one practice (Louisville) that CMS terminated for cause after completing the first three years. In Year 6, we excluded the two practices that left the

demonstration after Year 5 (Austin and Cleveland). In Year 7, we excluded the two practices that left the demonstration after Year 6 (Boston and Durham).

In Chapter 2 of this appendix, we begin by describing how the IAH practices operate and the characteristics of their patients. In Chapter 3, we describe how we identified the IAH group to evaluate the effect of the demonstration. In Chapter 4 of this appendix, we describe how we selected the comparison group. We then present the sources of data and measures for our quantitative analyses in Chapter 5. In Chapters 6 and 7 of this appendix, we describe the estimation of demonstration effects. Next, we present the methods and sources of data for our qualitative analysis in Chapter 8. Finally, in the last chapter, we discuss differences between incentive payment and evaluation results.

Exhibit A.1. IAH demonstration practices and number of beneficiaries by year

Demonstration practice location	IAH Year 1 Jun 2012–May 2013	IAH Year 2 Jun 2013–May 2014	IAH Year 3 Jun 2014–May 2015	IAH Year 4 Oct 2015–Sept 2016	IAH Year 5 Oct 2016–Sept 2017	IAH Year 6 Jan 2019–Dec 2019	IAH Year 7 Jan 2020–Dec 2020
Practices that participated in Years 1 to 7							
Brooklyn, New York	371	410	505	1,055	991	491	558
Dallas, Texas[b]	1,373	993	994	1,344	1,264	1,290	1,121
Flint, Michigan[b]	1,542	969	991	1,607	1,641	1,415	1,181
Jacksonville, Florida[b]	780	654	497	504	874	621	499
Lansing, Michigan[b]	524	526	702	652	611	608	458
Long Island, New York[a]	246	220	220	235	288	331	262
Milwaukee, Wisconsin[b]	514	553	634	575	489	450	394
Portland, Oregon	161	144	138	171	180	159	135
Richmond, Virginia (3 practices)[a,c,d]	290	311	280	277	323	310	253
Wilmington, Delaware[a]	225	254	241	213	235	232	197
Total IAH beneficiaries in Year 7 analyses	**6,026**	**5,034**	**5,202**	**6,633**	**6,986**	**5,907**	**5,058**
Practices that left the demonstration after Year 6[e]							
Boston, Massachusetts[a]	183	166	157	149	136	107	n.a.
Durham, North Carolina	828	1,066	1,267	1,705	1,974	1,979	n.a.
Practices that left the demonstration after Year 5[f]							
Austin, Texas[e]	911	684	601	686	574	n.a.	n.a.
Cleveland, Ohio[a,e]	268	316	337	331	378	n.a.	n.a.

Exhibit A.1 *(continued)*

Demonstration practice location	IAH Year 1 Jun 2012–May 2013	IAH Year 2 Jun 2013–May 2014	IAH Year 3 Jun 2014–May 2015	IAH Year 4 Oct 2015–Sept 2016	IAH Year 5 Oct 2016–Sept 2017	IAH Year 6 Jan 2019–Dec 2019	IAH Year 7 Jan 2020–Dec 2020
Practices that left the demonstration before Year 4[g]							
Atlanta, Georgia (2 practices)[c]	60	n.a.	n.a.	n.a.	n.a.	n.a.	n.a.
Chicago, Illinois (7 practices)[c,d]	202	n.a.	n.a.	n.a.	n.a.	n.a.	n.a.
Louisville, Kentucky	1,698	2,264	2,647	n.a.	n.a.	n.a.	n.a.
Stuart, Florida (2 practices)[c,d]	356	n.a.	n.a.	n.a.	n.a.	n.a.	n.a.

Source: Mathematica's analysis of data from the IAH implementation contractor and Medicare claims and enrollment data from the Chronic Conditions Warehouse.

[a] These practices participated in health systems affiliated with a university or medical school.

[b] These practices participated as part of the Visiting Physicians Association.

[c] Practices located in Richmond, Atlanta, Chicago, and Stuart participated as consortia.

[d] Richmond, Chicago, and Stuart started Years 1 to 3 on September 1.

[e] We exclude practices that chose to leave the demonstration after Year 6 from analyses of Year 7 in this report. We include these practices in analyses of Year 6.

[f] We exclude practices that left the demonstration after Year 5 because they could no longer meet the demonstration requirements from analyses of Year 6 and 7 in this report. We include these practices in analyses of Years 1 to 5.

[g] All analyses in this report exclude practices that left the demonstration before Year 5 because they could no longer meet the demonstration requirements. Atlanta, Chicago, and Stuart left the demonstration during Year 2. Louisville left the demonstration after Year 3.

n.a. = not applicable.

2. Description of IAH practices and beneficiaries

To understand the features of the IAH practices and identify the changes they made to improve care, we collected and analyzed interview data from the practices and analyzed their claims data. Every IAH site had substantial experience providing home-based primary care before the demonstration. The sites differed substantially, however, in their approaches to care, such as who was included on the care team; whether they were notified automatically of patients' hospital admissions or emergency department (ED) visits; whether they focused on serving in private homes or assisted living facilities; and whether they used a formal risk-stratification system, which groups the beneficiaries into high- and low-risk groups to aid in care planning. In this chapter, we summarize care delivery patterns according to each of the three types of practices: (1) Visiting Physicians Association (VPA) practices, (2) academic medical center practices, and (3) independent practices. We obtained information about the settings from which the practices provided care from claims data in Years 2 and 4 of the demonstration. We obtained other information from the IAH practices.[1] Exhibits A.2, A.3, and A.4 provide site-by-site information on practices' structural and operational characteristics and care delivery for practices that participated in Year 7 of the demonstration.[2]

Exhibits A.5 and A.6 provide site-by-site information from Medicare claims data on the percentage of visits by various types of providers and by mode of visit (in-person, telehealth, telephone, or online).[3] In this report, we refer to telehealth visits as those that include real-time audio and video communication between the clinician and the patient. Telephone visits include only real-time audio. Online visits reflect patient-initiated communication with the clinician using an online patient portal. For online visits, communications can occur over a 7-day period.

2.1. VPA practices

The five VPA practices (Dallas, Flint, Jacksonville, Lansing, and Milwaukee) had similar structural and operational characteristics (Exhibits A.2, A.3, and A.4). VPA is a for-profit corporation with multiple home-based primary care practices operating in several states; five of those practices participated in the demonstration. In each of these practices, most clinicians were physicians. As of Year 7, all of the practices

[1] Information in this chapter comes from site visits we conducted from February to May 2013 and February to July 2014. In January and February 2017 and in November 2019 to February 2020, we interviewed practices by telephone to confirm and update information.

[2] Information about the two practices that left the demonstration after Year 5, Austin and Cleveland, is available in the evaluation report covering Years 1 to 4 of the IAH demonstration and the evaluation report covering Years 1 to 5 of the IAH demonstration.

[3] We did not examine location of ambulatory visits—that is, home versus office visits—because all IAH practices provide home-based primary care nearly exclusively. For more information, see the evaluation report covering Years 1 to 4 of the IAH demonstration.

employed NPs, though more than half of visits were provided by a primary care physician (Exhibit A.5). None employed a social worker. In Year 7—the first year of the COVID-19 pandemic—three VPA practices provided about half of their visits in person, while the other two provided more than 70 percent in person (Exhibit A.6). All five practices provided visits via telehealth much more frequently than via telephone.

Historically, each practice had a patient care coordinator who was the main point of contact for patients and had access to the VPA corporate infrastructure for finance, human resources, data analytics, and data support. In Year 6, the VPA practices renamed the patient care coordinator position to nurse navigator. The nurse navigator role encompassed all patient care coordinator responsibilities and included managing care for patients with the highest hospital and ED use. On rare occasions when a clinician perceived a patient as needing extra support after discharge from the hospital or ED, a nurse navigator made a home visit.

Patients (both IAH beneficiaries and others) were assigned to a mobile care team consisting of one physician and one medical assistant.[4] VPA nurse navigators often visited the homes of their patients, although those visits were not billable. In four of the VPA sites, at least two-thirds of visits occurred in private homes. In Milwaukee, about two-thirds of visits occurred in assisted living or other group living facilities. None of the VPA practices conducted visits in other settings such as hospitals or skilled nursing facilities (SNFs).

In Year 6, the VPA practices reported continuing to foster relationships with SNFs and nursing homes to help coordinate care. In addition, to continue to strengthen existing relationships, clinicians frequently reached out to these care partners to remind them to coordinate with the practice when caring for their IAH patients.

Each VPA practice risk-stratified patients on the basis of their history of hospital admission and ED visits to determine the needed level of care and the frequency of proactive phone calls to patients and caregivers. Two practices developed relationships with hospitals and their staff; those staff notified the practice directly when one of its patients was hospitalized or visited the ED, whereas the remaining three practices received automated notices from hospitals.

2.2. Academic medical center practices

In Year 7, three IAH participants were part of nonprofit academic medical centers or health systems with academic missions: Long Island, Wilmington, and the three members of the Richmond-based consortium (Philadelphia, Richmond, and

[4] The term patients in this chapter refers to all patients treated by the practice regardless of IAH enrollment status or whether the beneficiary was in Medicare fee-for-service.

Washington).[5] The practice in Boston left the demonstration after Year 6 ended. In this section, we present qualitative data for each component of the consortium separately. This status gave these practices access to institutional resources and information technology systems and support. Clinicians in these settings were typically responsible for training and education in addition to clinical care, so many saw patients only part time.

In Long Island, physicians conducted most visits; in Wilmington, Philadelphia, Richmond, and Washington, NPs conducted most of the visits during Year 7 (Exhibits A.2 and A.5). Social workers were key members of the care team for all academic medical center practices because they coordinated home health services and referred patients to social services and supports. All but one academic medical center provided nonbillable visits, such as those conducted by social workers or nurses not acting under a physician's direction or as part of a home health episode. All academic medical center practices conducted most visits in private home settings; three (Long Island, Philadelphia, and Washington) conducted no visits in assisted living facilities (Exhibit A.3). Four of the five academic medical center practices conducted visits in other settings such as hospitals or SNFs. In Year 6, Wilmington reported working to strengthen relationships with local assisted living facilities via proactive outreach to ensure more coordinated care for their beneficiaries.

In Year 6, the Long Island and Washington practices reported testing e-consults with specialists to help manage patients' conditions and reduce Medicare expenditures. Clinicians sent questions and received advice by email from a specialist (for example, cardiologists or pulmonologists) to help manage homebound patients who could not travel to a specialist for an office visit. These two practices also used psychiatry consult services for clinicians to ask questions about managing patients' psychiatric conditions.

Academic medical centers varied in their use of technologies to facilitate care delivery and planning. Three of the five practices relied on clinical judgment to determine the level of care rather than using a formal risk-stratification system. Nearly all were notified automatically of patients' hospital admissions or ED visits from at least some hospitals with which they built relationships. In Year 7, all of the academic medical center practices provided a majority of visits in person (Exhibit A.6). Among visits that did not occur in person, the mode used most often by each practice differed. Long Island provided many more visits by telehealth than by telephone, while the Richmond-based consortium provided many more visits by telephone. Wilmington used telehealth and telephone equally.

[5] Three practices (Philadelphia, Richmond, and Washington) participated as one consortium, which the demonstration considers as one site for the purpose of calculating incentive payments.

2.3. Independent practices

In Year 7, the demonstration included two practices that began the demonstration as independent practices, Brooklyn and Portland; these practices differed in size, structure, and operations (Exhibits A.2–A.6). The practice in Durham dropped out of the demonstration late in the year (fall 2019). During the demonstration, the Brooklyn and Portland practices experienced the following changes related to ownership arrangements and partnerships:

- During Year 6, the independent practice in Brooklyn was acquired by Heal, a home-visiting service company that newly entered the Brooklyn market. In Year 6, the practice continued to participate in IAH with minimal modifications to the way it delivered care before being acquired by Heal; respondents from Heal reported treating all patients, regardless of IAH status, the same.
- HouseCall Providers in Portland signed a partnership agreement with CareOregon, a nonprofit health plan that provides care to Medicaid and Medicare beneficiaries, in May 2017 (during Year 5). Through this partnership, the practice aimed to increase access to home-based medicine across the Portland area and increase the share of their IAH-eligible patients who were dually eligible. Indeed, the percentage of IAH beneficiaries who were dually eligible increased from 19 percent in Year 5 to 27 percent in Year 6.

Although Brooklyn and Portland experienced organizational changes since we last collected qualitative data in April 2017, both practices reported during Year 6 that their care delivery for IAH beneficiaries remained stable in most respects. As in earlier demonstration years, Portland reported having staff dedicated to coordinating care for patients, but Brooklyn did not. In Portland, NPs continued to provide most of the visits as in earlier years, whereas in Brooklyn, the share of visits provided by NPs increased substantially during the demonstration (a change that began before the change in ownership) (Exhibit A.5). The sites of care for independent practices varied; Brooklyn conducted most visits in private home settings, and Portland conducted most visits in group living facilities (Exhibit A.3). Neither of the independent practices conducted visits in other settings such as hospitals or SNFs. Portland provided nonbillable visits to Medicare fee-for-service (FFS) beneficiaries by social workers and nurse care managers. Portland reported risk-stratifying patients as a way to determine the intensity of care the practice would provide, and Brooklyn reported relying on clinicians' judgment for these determinations. Both practices received automated notice of patient hospital admissions and ED visits from at least some hospitals. In the first year of the COVID-19 pandemic, Brooklyn provided nearly three-quarters of its visits in person; by contrast, Portland provided less than half in person and about 40 percent by telehealth (Exhibit A.6).

Exhibit A.2. IAH practices' structural characteristics, as of 2019

Site	Affiliation	Ownership	Participate in ACO (years)	Accept Medicare Advantage plans	Clinicians making home visits			Other staff involved in care team				
					Physicians	PAs	NPs	Care coordinators[a]	RNs	MAs	SWs	Others
VPA practices												
Dallas, TX	U.S. Medical Management	For profit	Yes (3)	No	13 FT	-	7 FT	1	2	31	-	
Flint, MI	U.S. Medical Management	For profit	Yes (3)	Yes	23 FT[b]	-	-	1	-	24	-	
Jacksonville, FL	U.S. Medical Management	For profit	Yes (3)	No	3 FT 2 PT	-	1 FT 3 PT	4	1	9	-	
Lansing/Ann Arbor, MI	U.S. Medical Management	For profit	Yes (3)	Yes	9 FT	1 FT	-	1	-	11	-	
Milwaukee, WI	U.S. Medical Management	For profit	Yes (3)	Yes	8 FT	-	2 FT	4	1	11	-	
Academic medical centers												
Long Island, NY	Northwell Health	Nonprofit	No	Yes	4 FT 4 PT	-	3 FT	-	5	-	5	
Philadelphia, PA	University of Pennsylvania	Nonprofit	No	Yes	2 PT	-	1 FT 1 PT	1	-	-	1	-
Richmond, VA	Virginia Commonwealth University	Nonprofit	No	Yes	2 FT	-	3 FT 6 PT	-	2	-	3	1 consulting pharmacist
Washington, DC	MedStar Health	Nonprofit	Yes (3)	Yes	6 FT	-	5 FT	5	1	-	5	1 LPN
Wilmington, DE	Christiana Care Health Systems	Nonprofit	Yes (5)	Yes	2 FT 4 PT	3 FT 1 PT	4 FT 3 PT	-	5	3	4	

Exhibit A.2 (*continued*)

Site	Affiliation	Ownership	Participate in ACO (years)	Accept Medicare Advantage plans	Clinicians making home visits			Other staff involved in care team				
					Physicians	PAs	NPs	Care coordinators[a]	RNs	MAs	SWs	Others
Independent practices												
Brooklyn, NY	Heal[c]	For profit	Yes (5)	Yes	3 FT	7 FT	21 FT	-	-	-	-	1 quality assurance nurse
Portland, OR	CareOregon[d]	Nonprofit	No	Yes	3 PT	2 PT	5 FT 11 PT	6	5	-	4	1 LPN, 1 pharmacist

Source: Information gathered from structured data collection before interviews with practice staff conducted from November 2019 to February 2020 (late in Year 6 of the demonstration to the beginning of Year 7).

Notes: For information on the two practices that left the demonstration after Year 5 (Austin and Cleveland), refer to the evaluation report covering Years 1 to 5 of the IAH demonstration. For information on the two practices that left the demonstration after Year 6 (Boston and Durham), refer to the evaluation report covering Years 1 to 6 of the IAH demonstration.

[a] Care coordinators are health professionals that help to manage a patient's care by monitoring and coordinating patients' care plans, connecting them with health care providers, and making telephone check-in calls. IAH sites use differing titles for this category of care, including nurse navigators, patient care coordinators, and care managers. For IAH practices, these staff generally are MAs, RNs, or LPNs.

[b] The Flint site did not provide a breakdown of physicians, PAs, or NPs.

[c] The Brooklyn site began the demonstration as an independent practice and was acquired by Heal during Year 6 (2019).

[d] The Portland site began the demonstration as an independent practice and signed a partnership agreement with CareOregon, a nonprofit health plan, in Year 5 (May 2017).

ACO = accountable care organization; FT = full-time; LPN = licensed practical nurse; MA = medical assistant; NP = nurse practitioner; PA = physician assistant; PT = part-time; RN = registered nurse; SW = social worker; VPA = Visiting Physicians Association.

Exhibit A.3. IAH practices' operational characteristics, as of 2019

Site	Visits per clinician per day	Clinicians' panel size	Nonbillable visits	Weekend visits	After-hours visits[a]	Share of visits in private residence[b]	Share of visits in ALF[b,c]	Visits outside of home
VPA practices								
Dallas, TX	9 to 10	175	Yes: nurse navigator	Occasionally[d]	No	88.3	11.1	No
Flint, MI	8 to 9	175	Yes: nurse navigator	Yes	No	77.6	22.4	No
Jacksonville, FL	8 to 9	175	Yes: nurse navigator	Occasionally[d]	No	68.0	32.0	No
Lansing/ Ann Arbor, MI	8 to 9	175	Yes: nurse navigator and home health company	Occasionally[d]	No	65.5	34.3	No
Milwaukee, WI	8 to 9	175	Yes: nurse navigator	Yes	No	27.9	70.6	No
Academic medical centers								
Long Island, NY	6	200	Yes: community paramedicine, RN, and SW	No	No	100.0	0.0	No
Philadelphia, PA	6	140	No	Yes	Yes: for urgent visits only; uncommon	94.8	0.0	Yes
Richmond, VA	3 to 4	40	Yes: nurse	No	No	96.1	0.4	Yes
Washington, DC	6	150	Yes: nurse	Yes	Yes: for regular visits; uncommon	86.6	0.0	Yes
Wilmington, DE	6	90 to 120	Yes: RN or SW	No	No	93.5	5.0	Yes
Independent medical practices								
Brooklyn, NY	8 to 10	120 to 130	Yes: visits to uninsured patients	Yes	Yes: for urgent and regular visits; common	99.8	0.0	No
Portland, OR	4	122	Yes: RN, social worker, or chaplain	Yes	No	16.0	83.9	No

Exhibit A.3 (*continued*)

Source: Information gathered from structured data collection before interviews with practice staff conducted from November 2019 to February 2020 (late in Year 6 of the demonstration to the beginning of Year 7).

Notes: For information on the two practices that left the demonstration after Year 5 (Austin and Cleveland), refer to the evaluation report covering Years 1 to 5 of the IAH demonstration. For information on the two practices that left the demonstration after Year 6 (Boston and Durham), refer to the evaluation report covering Years 1 to 6 of the IAH demonstration.

[a] After-hour visits are those done outside of the practice's normal business hours. This can vary from practice to practice.

[b] Percentage of visits in each location reflect data from Year 4. Visits may not sum to 100 percent because of claims that reflected care provided outside the home.

[c] ALF includes group homes and custodial care facilities.

[d] The term occasionally regarding weekend visits varies from practice to practice. The Dallas site defines it as "Saturdays occasionally." The Jacksonville site defines it as "a case-by-case basis up to 6/7 times a year." The Lansing/Ann Arbor site defines it as "up to the providers' discretion."

ALF = assisted living facility; RN = registered nurse; SW = social worker; VPA = Visiting Physicians Association.

Exhibit A.4. IAH practices' care delivery processes, as of 2019

Site	Formal risk-stratification classification	Remote access to patient's record, remote data collection, remote submission of orders	Notification of hospital admission or ED visit	Proactive outreach to patients or caregivers
VPA practices				
Dallas, TX	Yes. It is based on hospital or ED admissions: if patient has two or more visits in 60-day period, the patient is enrolled in an intensive care management program	Yes	Rely on hospital staff to notify practice	Yes. Call as needed based on acuity of patient
Flint, MI	Yes. It is based on hospital or ED admissions: if patient has two or more visits in 60-day period, the patient is enrolled in an intensive care management program	Yes	Automated notice from all sites within the state through HIE	Yes. Call as needed based on acuity of patient
Jacksonville, FL	Yes. It is based on hospital or ED admissions: if patient has two or more visits in 60-day period, the patient is enrolled in an intensive care management program	Yes	Automated notice from all sites within the state through HIE	Yes. Call as needed based on acuity of patient
Lansing/Ann Arbor, MI	Yes. It is based on hospital or ED admissions: if patient has two or more visits in 60-day period, the patient is enrolled in an intensive care management program	Yes	Automated notice from all sites within the state through HIE	Yes. Call as needed based on acuity of patient
Milwaukee, WI	Yes. It is based on hospital or ED admissions: if patient has two or more visits in 60-day period, the patient is enrolled in an intensive care management program	Yes	Rely on hospital staff to notify practice	Yes. Call as needed based on acuity of patient

Exhibit A.4 (*continued*)

Site	Formal risk-stratification classification	Remote access to patient's record, remote data collection, remote submission of orders	Notification of hospital admission or ED visit	Proactive outreach to patients or caregivers
Academic medical centers				
Long Island, NY	Yes. Determines level of proactive outreach and care team involved	No	Automated notices from some sites through EHR	Yes. Call as needed based on acuity of patient
Philadelphia, PA	No. Clinical judgment only	Yes	From within the health system, but not from other systems	Yes. Call as needed based on clinician's judgment
Richmond, VA	No. Clinical judgment only	Yes	Automated notice from practice's own hospital	No
Washington, DC	No. Clinical judgment only	Yes	Automated notices from some sites through EHR	Yes. Monthly call
Wilmington, DE	Yes. Software assesses patients and assigns level of acuity score, which determines level of proactive outreach and care team involvement	Yes	Automated notice from all sites within the state through HIE	Yes. Call as needed based on the acuity of the patient
Independent practices				
Brooklyn, NY	No. Clinical judgment only	Yes	Automated notice from some sites through EHR	Yes. Proactive outreach regarding vaccines in partnership with Merck
Portland, OR	Yes. Not fully rolled out, but it covered about three-quarters of patients in late 2019	Yes	Automated notice from all sites within the state through HIE	Yes. Call as needed based on acuity of patient and if patient was recently hospitalized

Exhibit A.4 (*continued*)

Source: Information gathered from structured data collection before interviews with practice staff conducted from November 2019 to February 2020 (late in Year 6 of the demonstration to the beginning of Year 7).

Notes: For information on the two practices that left the demonstration after Year 5 (Austin and Cleveland), refer to the evaluation report covering Years 1 to 5 of the IAH demonstration. For information on the two practices that left the demonstration after Year 6 (Boston and Durham), refer to the evaluation report covering Years 1 to 6 of the IAH demonstration.

ED = emergency department; EHR = electronic health record; HIE = health information exchange; VPA = Visiting Physicians Association.

Exhibit A.5. Percentage of visits from primary care physicians, nurse practitioners, and physician assistants at IAH practices for Years 2 and 7 among IAH beneficiaries

Site	Primary care physicians			Nurse practitioners			Physician assistants		
	Year 2	Year 7	Change	Year 2	Year 7	Change	Year 2	Year 7	Change
Mean per practice	69.1%	49.4%	-19.9	28.5%	43.5%	15.0	2.3%	7.1%	4.8
Academic health system									
Long Island	80.7%	66.2%	-14.5	19.3%	32.5%	13.2	0.0%	1.3%	1.3
Richmond[a]	32.0%	32.7%	0.7	68.0%	66.3%	-1.7	0.0%	1.0%	1.0
Wilmington	40.1%	32.3%	-7.8	43.6%	59.7%	16.1	16.3%	8.0%	-8.3
Independent									
Brooklyn	94.1%	25.6%	-68.5	5.9%	56.9%	51.0	0.0%	17.5%	17.5
Portland	5.8%	8.5%	2.7	87.1%	72.4%	-14.7	7.0%	19.0%	12.0
VPA									
Dallas	69.5%	60.0%	-9.5	30.5%	37.6%	7.1	0.0%	2.4%	2.4
Flint	98.0%	73.6%	-24.4	1.8%	24.8%	23.0	0.0%	1.6%	1.6
Jacksonville	92.2%	52.2%	-40.0	7.4%	45.2%	37.8	0.0%	2.6%	2.6
Lansing	91.1%	78.7%	-12.4	8.9%	7.9%	-1.0	0.0%	13.5%	13.5
Milwaukee	87.8%	64.3%	-23.5	12.2%	31.9%	19.7	0.0%	3.8%	3.8

Source: Mathematica's analysis of data from the IAH implementation contractor and Medicare claims and enrollment data from the Chronic Conditions Warehouse.

Notes: Results are not regression-adjusted. Visits reflect evaluation and management services, transitional care management services, annual wellness visits, advance care planning, and cognitive assessment and care plan services identified using procedure codes. Changes that were more than a 15 percentage point increase or 15 percentage point decrease since Year 2 are shaded in dark green and dark orange, respectively. Changes that were less than a 15 percentage point increase or 15 percentage point decrease since Year 2 are shaded in light green and light orange, respectively.

[a] Data reflect all three members of the Richmond-based consortium: Philadelphia, PA; Richmond, VA; and Washington, DC.

VPA = Visiting Physicians Association.

Exhibit A.6. Percentage of visits provided to IAH beneficiaries by IAH practices, by visit mode, Year 7

Site	In-person	Telehealth	Telephone	Online
Mean per practice	62.4	24.8	12.7	0.1
Academic health system				
Long Island	63.8	30.5	5.4	0.3
Richmond[a]	71.4	6.3	22.2	0.0
Wilmington	71.7	14.2	14.1	0.0
Independent				
Brooklyn	72.6	23.0	4.4	0.1
Portland	47.5	40.8	11.7	0.0
VPA				
Dallas	72.2	19.4	8.3	0.1
Flint	53.9	27.0	19.1	0.0
Jacksonville	70.7	20.4	8.9	0.0
Lansing	46.6	32.1	21.2	0.1
Milwaukee	54.1	34.0	11.9	0.0

Source: Mathematica's analysis of data from the IAH implementation contractor and Medicare claims and enrollment data from the Chronic Conditions Warehouse.

Notes: Results are not regression-adjusted. Visits reflect evaluation and management services, transitional care management services, annual wellness visits, advance care planning, and cognitive assessment and care plan services identified using procedure codes. Telehealth visits are those that include real-time audio and video communication between the clinician and the patient. Telephone visits include only real-time audio. Online visits reflect patient-initiated communication with the clinician using an online patient portal that takes place over a 7-day period.

[a] Data reflect all three members of the Richmond-based consortium: Philadelphia, PA; Richmond, VA; and Washington, DC.

VPA = Visiting Physicians Association.

3. Identifying the IAH beneficiaries

To comply with the legislation that established the IAH demonstration, the demonstration used a site-based enrollment process. Sites were responsible for ensuring that enrollees met health status and other clinical and programmatic requirements such as providing consent. The implementation contractor used administrative data and information provided by the sites to construct the list of enrolled beneficiaries as part of its work to calculate spending by IAH beneficiaries in each practice.

Although the implementation contractor used Medicare claims data, other administrative data, and information provided by the sites to construct the list of enrollees, Mathematica used only Medicare claims and other administrative data to identify the IAH group for the evaluation. (See Chapter 5 of this appendix for more information about the data sources we used to determine eligibility.) To measure the effect of the demonstration, we had to use the same data sources and approach to identify the IAH and comparison groups across all pre-demonstration and demonstration years. Information provided by the sites to construct the list of IAH enrollees was available for the demonstration years only, not the pre-demonstration years. In addition, no information other than administrative data was available for the comparison group. As a result, we used only administrative data to define the IAH group in each pre-demonstration and demonstration year rather than using the information the sites provided to the implementation contractor. We describe our process for defining the IAH group in this chapter. We describe our process for identifying the comparison group in Chapter 3 of this appendix.

The approaches of Mathematica and the implementation contractor to identifying eligible beneficiaries yielded different counts of IAH practices' beneficiaries in Years 1 to 7. After explaining these approaches in Sections 3.1 and 3.2, we provide details about reasons for differences in the counts of IAH practices' eligible beneficiaries in Section 3.3.

3.1. IAH implementation contractor's process for determining the sample of enrolled beneficiaries

The IAH sites identified beneficiaries they thought were eligible to participate in the demonstration; we list the eligibility requirements in Chapter 1 of the report. After providing these beneficiaries with information about the demonstration and visiting homes to explain it, the IAH sites enrolled willing participants and uploaded a list of potential enrollees to a reporting system created for the demonstration using a process established by the implementation contractor. The contractor then used administrative data to verify that each enrolled beneficiary had a qualifying hospital admission and used rehabilitation services in the previous 12 months, was covered by

Medicare Parts A and B, and was not enrolled in a Medicare Advantage plan as of the date of IAH enrollment.

In addition to verifying whether the beneficiaries enrolled by the practices had a qualifying hospital admission and used rehabilitation services, the implementation contractor also helped IAH sites identify potential beneficiaries for enrollment into the demonstration based on the eligibility criteria. In Year 7, the contractor identified beneficiaries who received at least one home visit from the demonstration practice and had qualifying hospital admission and rehabilitation service events, but whom the sites had not yet enrolled in the reporting systems; these beneficiaries were called potential enrollees. The implementation contractor began to include telephone (audio-only) visits and home visits made via telehealth in Year 7 to account for shifting patterns of care due to the COVID-19 pandemic and flexibilities offered by CMS during the public health emergency. The contractor provided the sites with information on the potential enrollees, and the sites then reviewed their records and assessed additional information about the beneficiaries' eligibility (such as whether they met the ADL and chronic condition criteria). Clinicians followed up with potential enrollees who met all demonstration criteria and enrolled them in the demonstration.

The implementation contractor set the enrollment date as the first day of the month after the beneficiary had a qualifying hospital admission, used rehabilitation services, and received a home visit by the IAH practice within the previous 12 months. The home/telephone visit by the practice might have occurred before or after the qualifying hospital admission and rehabilitation services as long as all three occurred within 12 months before the enrollment date.

If the beneficiary did not meet the demonstration eligibility criteria, the sites provided the implementation contractor with the reason for the beneficiary's ineligibility. Reasons sites reported for not enrolling beneficiaries whom the contractor identified as potential enrollees included the following: (1) the beneficiary did not meet the ADL or chronic condition criteria; (2) the beneficiary received primary care from another practice and the IAH practice was not considered the beneficiary's primary practice; (3) the beneficiary began receiving hospice care, moved into a nursing home, or died before receiving notification of his or her eligibility for the demonstration; and (4) the beneficiary refused to participate in the demonstration. If the IAH practice did not provide any reasons for ineligibility for a potential enrollee, the implementation contractor assumed that the beneficiary was eligible and added that person to the official demonstration enrollment records.

We refer to all beneficiaries confirmed as IAH participants in the implementation contractor's records as enrolled beneficiaries. Unless an IAH practice disenrolled a beneficiary—or a beneficiary died or was no longer enrolled in Medicare FFS—CMS allowed beneficiaries who enrolled in the demonstration in a given year to continue

in the demonstration whether or not they requalified in subsequent years as IAH eligible or had a home visit from the IAH practice in subsequent years.

3.2. Mathematica's process for identifying the sample of eligible and attributed beneficiaries for the evaluation

To identify beneficiaries eligible for the demonstration and attributed to a demonstration practice, Mathematica used different processes and data sources than those used by the implementation contractor and the IAH sites. As we explained earlier, our method for measuring the effect of the demonstration required us to use the same data sources and approach to identify the IAH and comparison groups across all pre-demonstration and demonstration years. We could not use enrollment in the demonstration as part of determining who would be in our sample because enrollment was based in part on information from the IAH practices. Therefore, the IAH group consisted of all beneficiaries eligible for the demonstration in that year according to our analysis of Medicare enrollment, claims, and assessment data.

We used the following criteria to identify beneficiaries for the IAH group:

- Enrollment in Medicare FFS
- Two or more ADLs that required human assistance
- Two or more chronic conditions
- Inpatient hospital admission or observation stay in the previous 12 months[7]
- Acute or subacute rehabilitation services use in the previous 12 months[8]
- Not in hospice or long-term care for the entire time they were eligible for the demonstration in a given year and not on hospice on the first day of demonstration eligibility

For beneficiaries enrolled in the demonstration, the eligibility date determined by Mathematica based on administrative data sometimes differed from the enrollment date determined by the implementation contractor. Mathematica set the eligibility date as the first day of the month following the last service use required to qualify for the demonstration. For example, if a beneficiary had a hospital admission in July 2020 and home health care in October 2020, that person would be eligible for demonstration Year 7 as of November 1, 2020.

In the following section, we explain why we identified the IAH group separately for each study year rather than retaining IAH beneficiaries in the sample until the demonstration ended. Then, we explain why we removed from the sample IAH practices that did not participate in the demonstration in a given year. Next, we

[7] This includes acute care, critical access, and psychiatric hospitals.

[8] This includes discharge from inpatient rehabilitation hospitals and rehabilitation units or SNFs and use of home health (but not necessarily discharge). We did not include long-term care hospitals.

describe how we used assessment data to measure limitations in ADLs. Finally, we explain how we used Medicare claims to attribute eligible beneficiaries to the IAH group.

3.2.1. Rationale for identifying the IAH group separately for each study year

An alternative to identifying the IAH group separately for each study year, which is called a repeated cross-sectional design, would be to retain beneficiaries in the IAH (and comparison) group until the demonstration ended, as long as the beneficiary remained enrolled in Medicare FFS. This approach is sometimes referred to as intent to treat, and it requires following all beneficiaries ever attributed to the IAH group during the demonstration, including those who did not requalify as eligible for IAH or stopped receiving care from the IAH practice.[9] This type of study design would have enabled us to evaluate outcomes for beneficiaries over a longer period, including for beneficiaries who did not requalify as eligible for IAH because they did not have another hospital admission or observation stay.

We did not use this type of study design for two primary reasons: (1) this approach can attenuate (that is, underestimate) true effects because it requires retaining beneficiaries in the intervention group who are no longer served by IAH practices, and (2) this approach would have required changing the length of the pre-demonstration period each time the demonstration was extended.

First, unless the intervention is expected to affect outcomes for several years after receipt of the intervention ended, the risk of underestimating true effects is high if a substantial minority of intervention group beneficiaries stopped receiving care from the IAH practice. This is the case with the IAH demonstration; among the beneficiaries from the Year 1 IAH group who remained alive and enrolled in FFS in Year 2, 38 percent were not attributed to the IAH practice in Year 2. In other words, 38 percent of beneficiaries from Year 1 did not receive care from an IAH practice in Year 2. We discuss attribution criteria in detail later in this section, but in general, attribution would have required the following: (1) at least one home visit from the IAH practice in Year 2, and (2) for a beneficiary who was alive more than 3 months in Year 2, a second visit from the IAH practice. The percentage of the intervention group who were no longer served by IAH practices would have increased over time as we retained beneficiaries in the sample for the entire demonstration (or at least for

[9] An intent to treat design retains beneficiaries in the analysis even if they stop receiving care from the IAH practice to avoid introducing bias related to attrition. That is, the intent to treat design considers attrition as part of the estimated population-level effect that should be included in the estimate. For example, if beneficiaries who continue to receive care from an IAH practice have lower expenditures than those who stop receiving care from an IAH practice, failing to include beneficiaries who no longer receive care from IAH practices could overstate the reduction in total expenditures for the full population compared to what might be expected, were the demonstration to be rolled out in another setting.

several years), diluting any actual effect of the IAH payment incentive on expenditures and other outcomes.

Second, a study design that retains beneficiaries in the sample for several years after they no longer meet IAH eligibility criteria carries a substantial risk of bias due to changes in health status (and resulting expenditures) that could have been associated with unobservable factors that led to IAH beneficiaries receiving care from IAH practices. For example, it is possible that declining cognitive status led some beneficiaries to start receiving home-based primary care from an IAH practice, and that (1) similar declines did not happen for the comparison group and (2) declining cognitive status tends to lead to higher (or lower) Medicare expenditures. If these unmeasured changes in cognitive status occurred over several years during the demonstration period—and not in the shorter pre-demonstration period—then retaining beneficiaries in the IAH sample for the entire demonstration could result in biased estimates of the effects of the demonstration. The way to reduce this risk would be to have a pre-demonstration period equal to the length of the demonstration period, as both periods would offer the same amount of time for beneficiaries to experience unmeasured changes. However, because the demonstration was extended four times by Congress, this would have required updating the pre-demonstration period multiple times—including generating new estimated effects for earlier demonstration years.

3.2.2. Rationale for removing from the sample IAH practices that did not participate in the demonstration in a given year

In all, 18 practices began the demonstration in 2012. Our estimates for Years 1 to 5 are based on data for the 14 practices that completed Year 5 of the demonstration, our estimates for Year 6 are based on data for the 12 practices that completed Year 6, and our estimates for Year 7 are based on data for the 10 practices that completed Year 7. An alternative to removing IAH practices from the sample if they did not participate in the demonstration in a given year would be to retain practices until the demonstration ended. This type of study design may have enhanced external generalizability of the impact estimates because the practices that left the demonstration may have been different from—and possibly less successful at reducing expenditures than—the practices that remained in the demonstration. We did not use this approach for the evaluation of the IAH demonstration for two reasons.

First, for a variety of reasons, we could not have included data from subsequent demonstration years for several of the practices that left the demonstration. For example, one practice that left the demonstration in Year 2 did not have enough home-based primary care patients in the pre-demonstration period to be included in our sample in any year. In other words, we could not include this practice in our analyses of the effect of the demonstration regardless of whether it continued

participating. Another practice was terminated from the demonstration by CMS for violating the Federal False Claims Act and subsequently closed. Including this practice in our analysis when it was in operation would have caused bias in our results when comparing expenditures for the IAH beneficiaries with those of a comparison group. A third practice—the practice that had a substantial influence on the estimated results for Year 5— stopped providing home-based primary care after it left the demonstration.[10] Even if we had wanted to include this practice until the demonstration ended, we could not have done so because we could not identify IAH beneficiaries for this practice after it stopped providing home-based primary care. We examined the relative influence of individual practices (see Chapter 6 of this appendix), however, and we considered this information when interpreting results.

Second, this demonstration was not designed to draw conclusions about the broad Medicare FFS population. The IAH practices were not selected to represent the national population of practices providing home-based primary care to Medicare beneficiaries with multiple chronic conditions and substantial functional limitations. Rather, among the pool of home-based primary care practices that volunteered for the demonstration, CMS selected 18 sites to represent different types of practices and geographic areas. Thus, even if we were to retain all IAH practices in the sample until the demonstration ended, we could not generalize the results of this study to Medicare FFS beneficiaries who received home-based primary care from practices other than those in the demonstration.

3.2.3. Eligibility and assessment data

We measured ADL limitations in accordance with the guidelines the IAH implementation contractor gave to IAH practices. Those guidelines stated that beneficiaries qualify as having an ADL limitation if they require any type of human assistance with the activity. The exception to this general guideline was for wheelchair use: use of a wheelchair as the primary mode of mobility with or without human assistance qualified as an ADL limitation for enrollment in the IAH demonstration.

To measure limitations in ADLs for the evaluation sample, we used assessment data from the given pre-demonstration or demonstration year. We used three sources of assessment data: (1) the Outcome and Assessment Information Set (OASIS), collected when beneficiaries receive home health care; (2) the Minimum Data Set (MDS), collected when beneficiaries receive SNF care; and (3) the Inpatient Rehabilitation Facility Patient Assessment Instrument (IRF-PAI), collected when beneficiaries receive inpatient rehabilitation facility care. All three data sets provided information about the extent to which beneficiaries could complete the six standard ADLs: dressing, bathing, toileting, transferring, ambulating, and feeding. Transferring includes transfer between bed and chair and excludes transferring to or from the bath or

[10] For more information, refer to the evaluation report covering Years 1 to 5 of the IAH demonstration.

toilet. Each assessment instrument has one or more data elements that indicate the extent of limitations, if any, for each of the six ADLs. If beneficiaries did not have any assessment data in a given year, they were ineligible for the demonstration in that year, and we did not include them in our sample.

We faced three challenges when measuring limitations with the six ADLs. First, each ADL is coded differently in each of the three data sets. Second, different providers collect ADL data at different points in time. Third, beneficiaries can have multiple assessments in a given year. Next, we discuss how we handled each of those three challenges.

Each ADL limitation is coded differently in each data set, and the codes do not always clearly define someone's need for human assistance to do the activity. We reviewed all of the values of each variable that measured ADL functioning. If the value for beneficiaries indicated that they required human assistance to do the activity safely, we classified them as requiring human assistance with that ADL.

In cases in which the level of functioning did not make clear that beneficiaries required human assistance to complete the activity, we erred on the side of not including patients. For example, one of the possible values for the transferring data element in an OASIS assessment was "able to transfer with minimal human assistance or with use of an assistive device," such as a walker. If beneficiaries had an OASIS assessment with that value for the transferring data element, we did not consider them to have a limitation that required human assistance for transferring based on that particular assessment. This conservative approach excluded from our sample beneficiaries who required a device but not human assistance, such as beneficiaries who could get out of bed alone when using a walker. It may also have excluded some people, however, who required human assistance and therefore could be IAH eligible.

Although we usually did not score beneficiaries as having a limitation if they required human assistance or an assistive device, we applied one exception to that rule. In accordance with the guidelines given to IAH practices by the implementation contractor, use of a wheelchair as the primary mode of mobility with or without human assistance qualified as an ADL limitation.

Different providers collect ADL data at different points in time. CMS requires that health care providers conduct OASIS, MDS, and IRF-PAI assessments at specific points in time. For example, beneficiaries who received skilled nursing services for a 60-day period may have had MDS data from assessments at admission, at discharge, and at the time of any significant changes in status. Because providers conduct each of these assessments at multiple points in time, we had to determine which assessments we would use in measuring ADL limitations to determine IAH eligibility. We used discharge assessments from all three data sets as well as interim assessments from the OASIS data set. We did not use admission or interim

assessments from the MDS and IRF-PAI because beneficiaries must be discharged from a SNF or inpatient rehabilitation facility before becoming eligible for IAH. Unlike with skilled nursing and inpatient rehabilitation services, beneficiaries can receive Medicare-funded home health care on the date they become eligible for IAH. Therefore, we included interim OASIS assessments in addition to discharge assessments to ensure we had the latest information in the study year.[11]

Beneficiaries can have multiple assessments in a given year. Beneficiaries could have had more than one assessment in a given year. For example, in one demonstration year, beneficiaries could have had three sets of assessment data: an interim OASIS assessment from home health care, a discharge OASIS assessment from home health care, and a discharge MDS assessment from skilled nursing care. When beneficiaries had more than one assessment in a given year, we kept the most recent assessment in which beneficiaries had at least two ADL limitations. We selected the most recent ADL assessment in which beneficiaries had at least two ADL limitations because we sought to identify beneficiaries who were least likely to recover from the ADL limitation. If beneficiaries had assessment data during a given year but not at least two ADL limitations in any of those assessments, they were ineligible for the demonstration in that year, and we did not include them in our sample. In addition, if beneficiaries did not have any assessment data in a given year, they were ineligible in that year, and we did not include them in our sample.

3.2.4. Attribution and enrollment data

In addition to determining eligibility for the demonstration, in each year we applied the following criteria for attributing a patient to a demonstration site (we used Medicare claims data for visits to the IAH practice that occurred between the date of eligibility for the demonstration and the end of the demonstration year):

- Residence in the same state as the demonstration practice.
- At least one evaluation and management or other home visit from the demonstration practice; home included private homes, assisted living facilities, group homes, and custodial care facilities. In Year 7, we updated our home visit identification methodology to mirror changes made by the IAH implementation contractor in response to flexibilities offered by CMS during the COVID-19 public health emergency. Specifically, we counted telephone (audio-only) evaluation and management visits and home visits provided via telehealth.
- For beneficiaries eligible for the demonstration for more than three months, at least one additional visit from the demonstration practice in the home, an assisted living facility, an office, or by telephone.

[11] Interim home health (OASIS) assessments do not include scoring on one activity: feeding. Because this item's effect on overall eligibility determination is small, we did not apply any adjustments to interim assessments.

The demonstration rules required that all patients of the IAH practice eligible for the demonstration be enrolled in the demonstration. Therefore, we required only one home visit for attribution to the IAH practice for beneficiaries eligible for the demonstration for three months or less. Some beneficiaries eligible for the demonstration for many months in a given year may have had only one visit with the IAH practice before returning to office-based primary care. To reduce the chance that the analysis sample would include beneficiaries who received only a single visit from the IAH practice, we required at least one additional visit from the practice for beneficiaries eligible for the demonstration for more than three months.

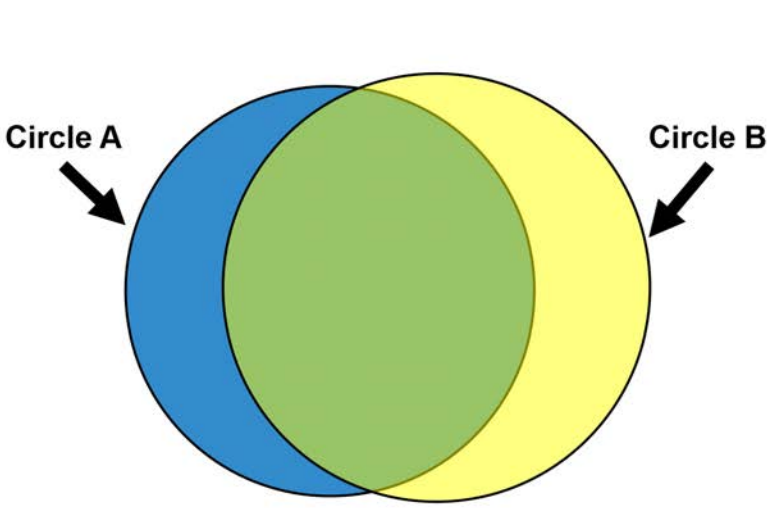
In each of the pre-demonstration and demonstration years, we refer to the beneficiaries who met eligibility criteria for IAH in administrative data and were attributed to a demonstration site as Mathematica-eligible IAH beneficiaries (or simply IAH beneficiaries). IAH beneficiaries were the treatment group for the evaluation. For beneficiaries to be in the IAH group for the evaluation, they had to meet the eligibility and attribution criteria outlined above according to Mathematica's analysis of Medicare enrollment, claims, and assessment data.

A beneficiary's enrollment (or non-enrollment) in the demonstration did not affect whether that person was in the IAH group for the evaluation. As we described, demonstration enrollment was based in part on data from the IAH practices, such as ADL limitations, chronic conditions, and residence in a long-term nursing home. In contrast, we excluded beneficiaries from the evaluation IAH group who were not eligible for the IAH demonstration and attributed to the IAH site according to administrative data (in Exhibit A.7, this is the part of Circle A excluding Circle B, or the blue crescent). We excluded those beneficiaries from the IAH group for two reasons: (1) we needed to identify the IAH group consistently in all study years, but demonstration enrollment data existed for the demonstration years only, not the pre-demonstration years; and (2) we could not replicate the enrollment process for comparison group members. In other words, we had no practice-reported data for identifying IAH beneficiaries in the pre-demonstration years, nor did we have such data for comparison group members in any year. Because our study design required that we use the same data sources to identify IAH and comparison beneficiaries in all years, we could not use practice-reported data to identify IAH beneficiaries in the demonstration years.

As shown in Exhibit A.7 and in the rest of this appendix, we use the term green oval to refer to beneficiaries who were enrolled in the demonstration and met its eligibility and attribution criteria in administrative data, according to Mathematica's analysis of those data. We use the term yellow circle to refer to beneficiaries who met the eligibility and attribution criteria for the demonstration regardless of whether they were enrolled in the demonstration. The yellow circle is the group we refer to as IAH beneficiaries (the treatment group for the evaluation). Enrollees who were not in the

evaluation IAH group (the blue crescent) were those who were enrolled but not confirmed as eligible for the demonstration or attributed to the IAH site according to administrative data.

Exhibit A.7. Groups of IAH beneficiaries based on different identification processes



Key:

Circle A (blue circle) = beneficiaries who were on the implementation contractor's list of IAH enrollees (IAH enrollees)

Circle B (yellow circle) = beneficiaries who Mathematica identified as eligible and attributed to an IAH practice (IAH beneficiaries used for the evaluation)

Intersection of Circle A and Circle B (green oval) = IAH enrollees Mathematica identified as eligible and attributed to an IAH practice

Circle A without Circle B overlap (blue crescent) = IAH enrollees Mathematica did not identify as eligible and attributed to an IAH practice

Circle B without Circle A overlap (yellow crescent) = beneficiaries Mathematica identified as eligible and attributed to an IAH practice but who were not enrolled in IAH

Note: Figure not drawn to scale.

After we identified an IAH beneficiary, that beneficiary remained in the sample for the rest of the demonstration or pre-demonstration year unless the person died or left Medicare FFS. For example, if an IAH beneficiary became eligible for the demonstration in February 2020 (month 2 of Year 7) and moved out of the IAH practice's geographic area or entered long-term care in April 2019, we continued to follow that beneficiary through the end of the study year (December 31, 2020, for all practices in Year 7).

Demonstration Year 1 (June 2012 to May 2013).[12] Mathematica identified 8,216 beneficiaries who met the demonstration eligibility criteria and were attributed to the demonstration practice during the first year (Exhibit A.8). This group represented the IAH group in the first year of the demonstration. It included 4,530 beneficiaries who were enrolled in the IAH demonstration according to the implementation contractor (in Exhibit A.7, the intersection of Circles A and B—the green oval) and 3,686 beneficiaries not enrolled in the IAH demonstration in Year 1 (in Exhibit A.7, Circle B excluding Circle A—the yellow crescent). The analysis sample did not include the 2,405 beneficiaries whom the implementation contractor identified as enrollees but

[12] For sites that began the demonstration in June 2012, month 1 was June. For sites that began the demonstration in September 2012, month 1 was September.

whom we did not find eligible for the demonstration using administrative data (in Exhibit A.5, Circle A excluding Circle B—the blue crescent).

Exhibit A.8. Numbers of beneficiaries based on different identification processes

Demonstration year	Mathematica-eligible IAH beneficiaries[a]			IAH-enrolled only (blue crescent)
	Mathematica-eligible and IAH-enrolled (green oval)	Mathematica-eligible only (yellow crescent)	Total IAH group (all Mathematica-eligible regardless of enrollment, yellow circle)	
1	4,530	3,686	8,216	2,405
2	4,564	2,702	7,266	4,059
3	4,498	3,066	7,564	4,718
4	6,019	3,485	9,504	5,663
5	5,950	4,008	9,958	6,407
6	2,692[b]	5,301[b]	7,993	3,696[b]
7	2,335	2,723	5,058	4,101

Source: Mathematica's analysis of data from the IAH implementation contractor and Medicare claims and enrollment data from the Chronic Conditions Warehouse. Data for Years 1 to 5 reflect the 14 IAH practices that participated in Year 5. Data for Year 6 reflect the 12 IAH practices that participated in Year 6. Data for Year 7 reflect the 10 IAH practices that participated in Year 7.

[a] This corresponds to the yellow circle in Exhibit A.7, which encompasses all Mathematica-eligible IAH beneficiaries (that is, those who met the demonstration eligibility criteria and were attributed to the demonstration practice).

[b] The withdrawal of Durham in month 10 of Year 6 was the largest contributor to the decrease in the number of Mathematica-eligible and IAH-enrolled beneficiaries from Year 5 to Year 6 and the increase in the number of beneficiaries who were only Mathematica-eligible. We retained Durham in the evaluation for Year 6 because it participated in the demonstration for most of the year. As a result of this early withdrawal, the implementation contractor did not identify any Durham patients as IAH enrollees in Year 6, which meant that all IAH beneficiaries we identified for Durham were only Mathematica-eligible.

Demonstration Year 2 (June 2013 to May 2014).[13] In Year 2, Mathematica identified 7,266 beneficiaries who met the demonstration eligibility criteria and were attributed to the demonstration practice. This group represented the IAH group in the second year of the demonstration. Of these 7,266 IAH beneficiaries, 4,564 were enrolled in the IAH demonstration in Year 2 (in Exhibit A.7, the green oval), and 2,702 beneficiaries were not enrolled (the yellow crescent). As in Year 1, the analysis sample for the evaluation did not include the 4,059 beneficiaries who were enrolled in the demonstration in Year 2 but whom we did not find eligible for the demonstration using administrative data (the blue crescent).

Beneficiaries enrolled but not eligible or attributed according to Mathematica in Year 2 (in Exhibit A.7, the blue crescent in Year 2) included people who enrolled for the

[13] For sites that began the demonstration in June 2012, month 1 was June. For sites that began the demonstration in September 2012, month 1 was September.

first time in Year 2. They also included two groups of beneficiaries who initially enrolled in Year 1 and continued to be enrolled in Year 2: those who were eligible and attributed according to administrative data in Year 1 (that is, those included in the yellow circle in Year 1) and those not eligible according to administrative data in Year 1 (the blue crescent in Year 1). The enrollment process did not require someone who was enrolled in Year 1 to meet the qualifications for enrollment in Year 2.

The IAH group for the Year 2 analysis sample consisted of the 7,266 beneficiaries identified as eligible and attributed by Mathematica (in Exhibit A.7, the yellow circle). As we explained, our method for measuring the effect of the demonstration required us to use the same data sources and approach to identify the IAH and comparison groups across all pre-demonstration and demonstration years. When we identified the Year 2 IAH beneficiaries, we did not consider whether a beneficiary was in the IAH group, comparison group, or neither group in Year 1. Therefore, the Year 2 IAH group included beneficiaries who were in the analysis sample in Year 1 and requalified in Year 2 by meeting eligibility and attribution requirements, as well as people not in the analysis sample in Year 1. It did not include beneficiaries who were in the IAH group in Year 1 but did not requalify for the IAH group in Year 2 because they failed to meet eligibility or attribution requirements. Including beneficiaries who qualified for the IAH group in Year 1 but did not requalify in Year 2 would potentially bias our estimates of the effect of the demonstration in Year 2, because non-requalifying beneficiaries in Year 2 could differ from the IAH beneficiaries in Year 1 and the pre-demonstration years—all of whom were selected without regard to which beneficiaries were in the IAH group in the prior year.

Demonstration Year 3 (June 2014 to May 2015).[14] In Year 3, Mathematica identified 7,564 beneficiaries who met the demonstration eligibility criteria and were attributed to the demonstration practices. This group represented the IAH group in the third year of the demonstration. Of these 7,564 IAH beneficiaries, 4,498 were enrolled in the IAH demonstration in Year 3, and 3,066 were not enrolled. These 7,564 beneficiaries included people in the analysis sample in Years 1 or 2 and who requalified in Year 3 by meeting eligibility and attribution requirements, as well as people not in the analysis sample in either of the first two years. These beneficiaries could be new patients who met the eligibility criteria or patients who previously received care from the IAH practice and did not meet the eligibility criteria for the demonstration until Year 3.

As in demonstration Year 1, the analysis sample for the evaluation did not include the 4,718 beneficiaries who were on the implementation contractor's enrollment list in Year 3 but whom we did not find eligible for the demonstration using administrative data. Beneficiaries enrolled but not eligible according to Mathematica in Year 3

[14] For sites that began the demonstration in June 2012, month 1 was June. For sites that began the demonstration in September 2012, month 1 was September.

included those who enrolled for the first time in Year 3. Beneficiaries enrolled but not eligible according to Mathematica in Year 3 also included beneficiaries who initially enrolled in Years 1 or 2, continued to be enrolled in Year 3, but did not requalify for the demonstration in Year 3 because they failed to meet eligibility or attribution requirements.

Demonstration Year 4 (October 2015 to September 2016). In Year 4, Mathematica identified 9,504 beneficiaries who met the demonstration eligibility criteria and were attributed to the demonstration practices. This group represented the IAH group in the fourth year of the demonstration. Of these 9,504 IAH beneficiaries, 6,019 were enrolled in the demonstration, and 3,485 were not. These 9,504 beneficiaries included people who were in the analysis sample in Years 1, 2, or 3 and requalified in Year 4 by meeting eligibility and attribution requirements, as well as people not in the analysis sample in any of the first three years.

As in Years 1 to 3, the analysis sample for the evaluation did not include the 5,663 beneficiaries enrolled in the IAH demonstration in Year 4 but who we found ineligible for or attributed to the demonstration using administrative data. The Year 4 IAH group also did not include beneficiaries who initially enrolled in Years 1, 2, or 3; continued to be enrolled in Year 4; but did not requalify for the demonstration in Year 4 because they failed to meet eligibility or attribution requirements.

The IAH group was substantially larger in Year 4 than in previous demonstration years. For all sites combined, this group increased 26 percent from Year 3 to Year 4. Five sites had increases of more than 20 percent from Year 3 to Year 4: Brooklyn, Durham, Dallas, Flint, and Portland. This increase could reflect the expansion of existing IAH practices. Brooklyn merged with another home-based primary care practice, and the Durham practice has expanded throughout North Carolina since the demonstration began. In Year 4, Dallas expanded into a new geographic area, and Flint added clinicians in its existing geographic area. Finally, Portland's sample size in Year 4 was larger than in Year 3 but was about the same size as in Year 1. The increase from Year 3 to Year 4 also could have been caused in part by some IAH practices participating in accountable care organizations (ACOs) in Year 4. Several practices did so in Year 4, including three of the five practices with the largest increases in sample sizes: Brooklyn, Dallas, and Flint. Other providers in the ACO may have referred some patients to the IAH practice. We explore the implications of ACO participation in Chapter 7 of this appendix.

Demonstration Year 5 (October 2016 to September 2017). In Year 5, Mathematica identified 9,958 beneficiaries who met the demonstration eligibility criteria and were attributed to the demonstration practices. Of these 9,958 IAH beneficiaries, 5,950 were enrolled in the demonstration, and 4,008 were not. These 9,958 beneficiaries included people in the analysis sample in Years 1 to 4 who requalified in Year 5 by

meeting eligibility and attribution requirements, as well as those not in the analysis sample in any of the first four years.

As in Years 1 to 4, the analysis sample for the evaluation did not include the 6,407 beneficiaries who were enrolled in the IAH demonstration in Year 5 but whom we found ineligible for or attributed to the demonstration using administrative data. The Year 5 IAH group also did not include beneficiaries who initially enrolled in Years 1 to 4, continued to be enrolled in Year 5, but did not requalify for the demonstration in Year 5 because they failed to meet eligibility or attribution requirements.

The noticeable increase in size from Year 3 to Year 4, in which the IAH group increased by 26 percent, did not repeat in Year 5. Rather, the sample size increased by only 5 percent, consistent with the observed increase from Years 2 to 3 (4 percent). This stability suggests that the observed increase in the overall size of the IAH sample in Year 4 was likely the result of events that may not reoccur in subsequent years, such as Brooklyn merging with another home-based primary care practice. This finding would be consistent with the proposed reasons for sample size increases outlined in the Year 4 summary above.

Demonstration Year 6 (January to December 2019). In Year 6, Mathematica identified 7,993 beneficiaries who met the demonstration eligibility criteria and were attributed to the demonstration practices. Of these 7,993 IAH beneficiaries, 2,692 were enrolled in the demonstration, and 5,301 were not. These 7,993 beneficiaries included people in the analysis sample in Years 1 to 5 who requalified in Year 6 by meeting eligibility and attribution requirements, as well as those not in the analysis sample in any of the first five years.

As in Years 1 to 5, the analysis sample for the evaluation did not include the 3,696 beneficiaries who were enrolled in the IAH demonstration in Year 6 but whom we found ineligible for or attributed to the demonstration using administrative data.[15] The Year 6 IAH group also did not include beneficiaries who initially enrolled in Years 1 to 5, continued to be enrolled in Year 6, but did not requalify for the demonstration in Year 6 because they failed to meet eligibility or attribution requirements.

The number of beneficiaries in the IAH group according to our analysis of administrative data decreased by 20 percent from Year 5 to Year 6. This decrease was a departure from the increases seen in previous years. Nearly all of this decrease can be attributed to three factors: (1) Cleveland and Austin leaving the demonstration

[15] When reviewing claims data for IAH enrollees at Brooklyn in Year 6, we found a large increase in the number of beneficiaries who were IAH-eligible and received care from a Brooklyn clinician but who were not included in the evaluation sample because their claims did not have the necessary procedure code(s) for attribution to the IAH practice. There were 55 such beneficiaries in Year 5 and 280 such beneficiaries in Year 6. Most of the reviewed claims for these 280 beneficiaries were for home health recertification with patient not present, suggesting that the observed changes may have been the result of changes in coding processes coinciding with a change in ownership from Years 5 to 6.

after Year 5, (2) a substantial decrease in the number of IAH beneficiaries identified for Brooklyn, and (3) a drop in clinicians treating patients at the Jacksonville practice. If we were to exclude those four practices in both years, the number of IAH beneficiaries was nearly as large in Year 6 as in Year 5.

The number of beneficiaries who were in the IAH group according to our analysis of administrative data but not enrolled in the demonstration increased by 32 percent from Year 5 to Year 6, a departure from the trend in previous years (13 to 15 percent year-over-year increases from Year 3 to Year 5). Most of this change was a result of the withdrawal of Durham from the demonstration toward the end of Year 6. The implementation contractor finalizes its list of enrollees after each demonstration year, reflecting practices that completed participation in that year. Beneficiaries from Durham were not identified as enrollees because of the midyear withdrawal, which means that none of the IAH beneficiaries from Durham that we included in the evaluation in Year 6 were enrolled in the demonstration.

In addition to site-level variation from Year 5 to Year 6, there were two data artifacts that may have contributed to the year-over-year sample variability. First, at the time we identified our Year 6 sample, assessment data we use to measure assistance with ADLs were unavailable for some beneficiaries. OASIS data were available for all assessments administered from January 2019 to November 2019 but were missing for a small share of those administered in December 2019. IRF-PAI data were available for all assessments administered from January 2019 to September 2019 but were missing for a small share of those administered from October 2019 to December 2019.

Second, the IRF-PAI underwent a major revision between Year 5 and Year 6, and most IRF-PAI variables we used for ADL identification in previous study years were not available in Year 6. The two versions of the IRF-PAI administered in 2019 contained other items that evaluated ADLs. Using these other variables, we constructed new measures of ADLs requiring human assistance that are as similar as possible to the original measures. Changes in the IRF-PAI assessment, however, may have led to slight differences in ADL measures in Year 6 compared with earlier years.

Neither of these issues—missing data for OASIS and IRF-PAI and the change in the ADL measures in the IRF-PAI—poses a substantial risk of bias to the estimated impacts of the demonstration in Year 6. This is because these issues affect both the IAH and comparison groups and the data were missing for all beneficiaries during certain months (unrelated to beneficiary health status or provider). Also, only one month of OASIS data were missing, and only a small share of the sample in each year qualified as IAH-eligible based on IRF-PAI data.

Demonstration Year 7 (January to December 2020). In Year 7, Mathematica identified 5,058 beneficiaries who met the demonstration eligibility criteria and were

attributed to the demonstration practices. Of these 5,058 IAH beneficiaries, 2,335 were enrolled in the demonstration, and 2,723 were not. These 5,058 beneficiaries included people in the analysis sample in Years 1 to 6 who requalified in Year 7 by meeting eligibility and attribution requirements, as well as those not in the analysis sample in any of the first six years.

As in Years 1 to 6, the analysis sample for the evaluation did not include the 4,101 beneficiaries who were enrolled in the IAH demonstration in Year 7 but whom we found ineligible for or attributed to the demonstration using administrative data. The Year 7 IAH group also did not include beneficiaries who initially enrolled in Years 1 to 6, continued to be enrolled in Year 7, but did not requalify for the demonstration in Year 7 because they failed to meet eligibility or attribution requirements.

The number of beneficiaries we attributed to the IAH group using administrative data decreased by 37 percent from Year 6 to Year 7. Much of this decrease can be attributed to Boston and Durham leaving the demonstration after Year 6. Excluding those two practices, the decrease from Year 6 to Year 7 was 14 percent. There was a general decrease in the number of beneficiaries attributed to each site; other than Brooklyn, the number of IAH beneficiaries at each site decreased by between 12 and 25 percent. There was a 14 percent increase in the number of IAH beneficiaries at the Brooklyn practice in Year 7, which we describe in more detail in this section. Some of the factors that led to the general decrease in the number of beneficiaries attributed to most sites are decreasing enrollment in Medicare FFS (because of increasing enrollment in Medicare Advantage) and changes in the availability of the ADL assessment data we used to identify beneficiaries eligible for the demonstration (largely because of factors related to the COVID-19 pandemic), which we describe in more detail in this section.

The number of beneficiaries who were in the IAH group (according to our analysis of administrative data) but not enrolled in the demonstration decreased by 49 percent from Year 6 to Year 7, a departure from increases in prior years. This large decrease was caused by three primary factors. First, the number of beneficiaries who were in the IAH group for the evaluation but not enrolled in the demonstration in Year 6 was unusually high because Durham withdrew from the demonstration late in Year 6 (discussed in greater detail earlier in this chapter). If we compare Year 7 to Year 5, the number of beneficiaries who were in the IAH group (according to our analysis of administrative data) but not enrolled in the demonstration in Year 7 was only 5 percent higher than in Year 5.

Second, we included IAH-eligible patients cared for by 12 clinicians (9 in Brooklyn and 3 in Flint) who were not used by the IAH implementation contractor to identify the enrolled sample. These 12 clinicians were not included in the list of clinicians provided to us by the implementation contractor because sites did not identify them as IAH clinicians in the IAH reporting system; we identified them because they

provided visits to a number of IAH-eligible beneficiaries. We verified the status of these clinicians with the sites and added them to the clinician list we used to identify IAH beneficiaries for the evaluation. By including these clinicians, we captured 280 additional beneficiaries in the IAH group who were not in the enrolled sample. Most of these beneficiaries were attributed to the Brooklyn practice.

Third, there were changes in the availability of the ADL assessment data we used to identify beneficiaries eligible for the demonstration. From Year 6 to Year 7, we identified 10 percent fewer beneficiaries with two or more ADLs in administrative data. This reduction was caused by multiple factors:

- Because of issues linking Medicare beneficiary identifiers to assessment data, a small share of OASIS and IRF-PAI data were not available for assessments administered between January and May of 2020. This issue was a continuation of the data anomaly originally identified late in Year 6, which we discussed earlier in this chapter.
- The number of MDS assessments decreased by 26 percent from March to December 2020 relative to the same timeframe in Year 6. This decrease was due to (1) less SNF and nursing facility use during the pandemic in 2020 than in previous years and (2) CMS temporarily waiving MDS assessment timeframe requirements during the COVID-19 public health emergency for SNFs and allowing states to waive the timeframe requirements for nursing facilities if needed.[16] Concerns about the risk of contracting COVID-19 might have led to a decrease in the number of beneficiaries who received care at a SNF (because of a decrease in the share of beneficiaries who were admitted to a SNF after hospital discharge and fewer overall hospital stays). When CMS reinstated the MDS timeframe requirement, they noted that most facilities completed assessments in a timely manner.[17] This suggests that decreased SNF and nursing facility use during the first year of the pandemic was primarily responsible for the decrease in MDS assessments, rather than the timeframe waiver.

To assess how the change in ADL data might have affected sample identification, we examined the percentage of beneficiaries we identified as needing human assistance with two or more ADLs by assessment type (Exhibit A.9). For both the IAH and comparison groups, the percentage of beneficiaries identified by OASIS only increased from Year 6 to Year 7: an increase of 4.0 percentage points for IAH beneficiaries and 4.1 percentage points for comparison beneficiaries. At the same time, the percentage identified by MDS only decreased: a decrease of 1.2 percentage points for IAH beneficiaries and 3.2 percentage points for comparison beneficiaries. The missing assessment data probably do not pose a substantial risk of bias to the estimated impacts of the demonstration in Year 7. These issues affected both groups,

[16] https://www.cms.gov/files/document/covid19-emergency-declaration-health-care-providers-fact-sheet.pdf.

[17] https://www.cms.gov/files/document/qso-21-17-nh.pdf.

and differences between Years 6 and 7 in the share of beneficiaries identified by particular sources of assessment data were similar for the two groups.

Exhibit A.9. Percentage of beneficiaries identified as needing human assistance with two or more ADLs, by assessment data source

Assessment	IAH			Comparison[a]		
	Year 6[b]	Year 7	Percentage point difference	Year 6[b]	Year 7	Percentage point difference
OASIS only	56.4	60.4	4.0	38.3	42.4	4.1
MDS only	11.5	10.3	-1.2	34.9	31.7	-3.2
IRF-PAI only	0.6	0.6	0.0	1.8	2.6	0.8
MDS and IRF-PAI	0.5	0.4	-0.1	1.2	1.5	0.3
OASIS and MDS	27.7	24.1	-3.6	20.2	17.4	-2.8
OASIS and IRF-PAI	1.7	2.3	0.6	2.2	2.7	0.5
OASIS, MDS, and IRF-PAI	1.7	1.9	0.2	1.5	1.7	0.2

Source: Mathematica's analysis of data from the IAH implementation contractor and Medicare claims, assessment, and enrollment data from the Chronic Conditions Warehouse.

[a] See Chapter 4 of this appendix for details regarding the matched comparison beneficiary selection process.

[b] Year 6 percentages are restricted to sites that also participated in Year 7.

ADL = activities of daily living; IRF-PAI = Inpatient Rehabilitation Facility Patient Assessment Instrument; MDS = Minimum Data Set; OASIS = Outcome and Assessment Information Set.

These data make clear that, in Years 6 and 7, the share of beneficiaries identified only by OASIS was nearly 50 percent higher for the IAH group than the comparison group, while the share identified only by the MDS was two-thirds lower for the IAH group. This difference is not surprising given that, even before the demonstration, IAH-eligible patients of IAH practices were more likely to use home health at all and used home health for more days than comparison beneficiaries.[18] Given that OASIS data are generated when home health services are used while MDS are generated when SNF or nursing facility services are used, these differences suggest the possibility of unobservable differences between IAH and comparison beneficiaries—such as preferences for receiving health care in the home, caregiver support, and coordination between primary care and home health providers. The fact that three times as many comparison beneficiaries were identified using MDS data relative to the number of IAH beneficiaries could pose a risk of bias in Year 7. This would occur if unobserved factors associated with facility use (which led to the MDS assessment) or home health use (which led to the OASIS assessment) had a different relationship

[18] For more information about the probability of using home health and average number of home health days per beneficiary per year, refer to the evaluation report covering Years 1 to 4 of the IAH demonstration.

with expenditures and other outcomes during the COVID-19 pandemic (Year 7) than in earlier years.

3.3. Reasons for the differences between demonstration enrollment and evaluation analysis cohorts

The evaluation analysis group identified by Mathematica and the enrolled group identified by the implementation contractor differed for two overarching reasons: (1) the use of different data sources and (2) the use of different analytic techniques. The primary reason Mathematica used different data sources and analytic techniques was that the implementation contractor had to identify only an IAH group, whereas Mathematica had to identify both an IAH group and a comparison group. Because Mathematica had to use the same procedures to identify both groups, and we could not obtain clinical data from the comparison group's primary care providers, we relied on administrative data alone when identifying the IAH group for the evaluation. This approach was in contrast to that of the implementation contractor, which used administrative data and data from IAH practices to identify IAH enrollees. This difference resulted in Mathematica excluding some beneficiaries identified as enrollees by the implementation contractor and including some beneficiaries in the IAH group for the evaluation who were excluded by the implementation contractor. We presented a detailed discussion regarding the differences in the samples for the evaluation versus enrollment in a prior report; we highlight key findings in this chapter.[19]

3.3.1. Reasons some IAH enrollees did not meet Mathematica's eligibility or attribution criteria

The use of different data sources was the primary reason Mathematica excluded some beneficiaries from the IAH group that the implementation contractor identified as enrollees. In each demonstration year, most IAH enrollees not identified by Mathematica did not meet the ADL criterion because they had missing or insufficient ADL information in the assessment data. In contrast, the contractor used information provided by the IAH practices to determine whether a beneficiary required human assistance with at least two ADLs. In addition, the number of enrollees that Mathematica did not find eligible for the demonstration increased over time, because beneficiaries remained on the IAH enrollment list from one year to the next regardless of whether they met IAH eligibility criteria again.

[19] For more information, refer to the evaluation report covering Years 1 to 4 of the IAH demonstration.

3.3.2. Reasons some beneficiaries found eligible and attributed by Mathematica were not enrolled

Among those who were in the IAH group for the evaluation but were not IAH enrollees, Mathematica identified three groups of beneficiaries:

- **Beneficiaries not found to be eligible by the implementation contractor based on administrative data.** As part of determining eligibility for enrolling in the demonstration, the contractor considered the dates that the beneficiary had a hospital admission, used rehabilitation services, and had a home visit from the IAH practice. Mathematica, however, considered only the dates of the qualifying hospital admission and rehabilitation services stay. Mathematica did not rely on the date of a home visit when measuring the 12-month period and setting the demonstration eligibility date because we could not replicate that requirement for the comparison group, who did not receive home-based primary care and therefore received no home visit.
- **Beneficiaries excluded from enrollment based on information from IAH practices.** The reason sites offered most frequently for excluding a beneficiary from enrollment was that the beneficiary did not meet the ADL criterion. The implementation contractor used information provided by the IAH practices to determine whether a given ADL required human assistance, which provided a more nuanced picture of ADL severity. Mathematica used only administrative data when identifying ADLs that required human assistance because information from clinicians was not available for the comparison group.
- **Beneficiaries who disenrolled from the demonstration.** Enrollees may voluntarily disenroll from the demonstration when they change clinicians within the practice service area, are discharged by the practice, decline home care, or elect hospice and change clinicians. If the beneficiary voluntarily disenrolled within six months of enrollment in the demonstration, the implementation contractor did not identify that beneficiary as an enrollee in the final enrollment list for a given year. Mathematica did not exclude a beneficiary who voluntarily disenrolled within six months because we could apply no such restriction to the comparison group.

3.4. Characteristics of IAH beneficiaries

To understand how characteristics of IAH-eligible beneficiaries differed from the average Medicare beneficiary at the start of the IAH demonstration, we used Medicare administrative data to identify beneficiaries who met IAH eligibility criteria and were treated by IAH practices in the year before the demonstration. In the year before the demonstration, more than half of beneficiaries who met the IAH eligibility criteria data were age 80 or older, and 40 percent were dually eligible for Medicare and Medicaid (Exhibit A.10). The demonstration eligibility criteria focused on

Medicare beneficiaries who were chronically ill and disabled. As a result, about 43 percent of IAH-eligible beneficiaries had 10 or more chronic conditions, and 55 percent required human assistance with at least five ADLs. On average, IAH-eligible beneficiaries incurred nearly $4,400 in Medicare expenditures PBPM in the year before the demonstration. They had an average of 1.8 hospital admissions and 2.9 ED visits per year.

IAH-eligible beneficiaries were more likely to be dually eligible, to be older, to have more chronic conditions, and to have a higher death rate than the average Medicare beneficiary. Among the IAH states, the average percentage of Medicare beneficiaries who were dually eligible in 2013 was 19 percent (Kaiser Family Foundation). In 2012, 36 percent of beneficiaries who resided in the community (not a facility) were older than age 75, 26 percent had five or more chronic conditions, and 3 percent died over the course of the survey year (CMS 2012). IAH-eligible beneficiaries also struggled with daily activities at a higher rate than the average Medicare beneficiary. Only 12 percent of Medicare beneficiaries who resided in the community reported difficulties in performing three or more ADLs without human assistance or special equipment, such as a walker or grab bar (CMS 2012).

Exhibit A.10. Demographic characteristics and health status, Medicare expenditures, and service use of beneficiaries who were eligible for IAH and treated by IAH practices in the year before the demonstration

Variable name	Value for beneficiaries eligible for IAH in the year before the demonstration
Demographic characteristics and health status	
Percentage age 80 or older	51.7
Percentage dually eligible for Medicare and Medicaid	40.1
Average HCC score	3.5
Percentage with 10 or more chronic conditions	42.7
Percentage requiring human assistance with at least five activities of daily living	55.0
12-month death rate (percentage)	18.1
Average Medicare expenditures per beneficiary per month	
Total	$4,397
Inpatient hospital services	$1,741
Skilled nursing facility services	$605
Home health (Parts A and B)	$781
Hospice services	$153
Outpatient services	$253
Clinician/supplier services	$715
Durable medical equipment	$150

Exhibit A.10 (*continued*)

Variable name	Value for beneficiaries eligible for IAH in the year before the demonstration
Average numbers of key health service use events per beneficiary per year	
Number of hospital admissions[a]	1.8
Number of potentially avoidable hospital admissions[b]	0.5
Number of ED visits	2.9
Probability (as a percentage) of having a qualifying hospital discharge and unplanned readmission within 30 days of discharge	19.6

Source: Mathematica's analysis of data from the IAH implementation contractor and Medicare claims and enrollment data from the Chronic Conditions Warehouse for IAH practices that participated in Year 6.

[a] The number of hospital admissions includes observation stays.

[b] The number of potentially avoidable hospital admissions includes observation stays. A potentially avoidable hospital admission is one in which appropriate primary and specialty care may prevent or reduce the need for a hospital admission.

ED = emergency department; HCC = Hierarchical Condition Category.

4. Identifying the comparison group

In this chapter, we begin by describing how we used Medicare administrative data to identify a potential comparison group of beneficiaries who were eligible for the demonstration, lived in the same area as the IAH beneficiaries, and did not receive home-based primary care. Next, we present the methods and results of propensity score matching. Finally, we present the number of IAH and comparison beneficiaries and eligible months in the evaluation sample.

4.1. Identifying the potential comparison group

To identify the potential comparison group beneficiaries, we relied on administrative data. We identified a set of potential comparison beneficiaries from each state in each year. We based our analyses on data for two pre-demonstration years and seven demonstration years. Beneficiaries who had no visits to any of the demonstration practices in the study year and met all demonstration eligibility criteria were eligible to be in the potential comparison group for all sites in that state in that year. For example, a beneficiary who lived in Michigan, had no visits from any IAH practice, and met all demonstration eligibility criteria in Year 2 was in the potential comparison group for Flint and Lansing. We refer to these comparison groups as potential because we identified the final comparison groups using propensity score matching (described later in this chapter). Because we sought to compare beneficiaries who primarily received in-home physician care with those who did not receive such care, we excluded from the potential comparison group all beneficiaries who had two or more home visits from any clinician (in person or via telehealth) during or after their first month of eligibility through the end of the study year. In addition, we excluded all beneficiaries who had any visit from an IAH practice in the study year. As with the IAH beneficiaries, we did not assess whether potential comparison beneficiaries had home visits before the first month of eligibility.

In addition, to control for possible geographic variation in practice styles, access to services, and costs, we restricted our comparison groups to beneficiaries who lived in the ZIP codes served by the demonstration practices. The list of ZIP codes served by a demonstration practice in a given year reflected all ZIP codes in which the practice's IAH beneficiaries lived in that year, according to beneficiary address information in Medicare administrative data. For example, if a site operated in one state and had at least one IAH beneficiary who lived in each of 57 ZIP codes in that state during demonstration Year 1, the potential comparison group for that site in Year 1 included all beneficiaries who met demonstration eligibility requirements, had no visits to any demonstration practice in that year, had no more than one home-based primary care visit in that year, and lived in one of those ZIP codes. We used this ZIP code–based restriction for all practices in all years.

For the six practices located in states that had two demonstration practices (Brooklyn and Long Island in New York as well as Flint and Lansing in Michigan in Years 1 to 7; Austin and Dallas in Texas in Years 1 to 5), some ZIP codes contained IAH beneficiaries for two practices. We could not simply restrict potential comparison beneficiaries only to those living in the ZIP codes represented by beneficiaries served by the IAH practice in a given year, because it would have allowed a single potential comparison beneficiary to be selected as a matched comparison for two IAH beneficiaries in different practices. In those cases, we identified the potential comparison group by conducting a preliminary propensity score matching (using the same model to predict treatment status we describe below) to split the comparison sample into two potential comparison groups.[20] For each pair of sites located in the same state, we included in the preliminary model all of the IAH beneficiaries in those two sites as well as all beneficiaries in the comparison pool for both sites after applying the ZIP code restriction. Each comparison beneficiary was matched to an IAH beneficiary in one of the two sites; this process determined the site potential comparison pool to which the beneficiary was assigned. After using preliminary matching to split the overlapping comparison sample into two potential comparison groups (one group per practice), we matched IAH beneficiaries to the potential comparison group for each practice using the same approach as for other sites.

As with IAH beneficiaries, we again identified beneficiaries in the matched comparison group in earlier demonstration years as potential comparison beneficiaries if they met all IAH eligibility requirements in Year 7.

4.2. Propensity score matching methods

For each analysis year before and after the demonstration began, we used propensity score matching to create a comparison group of nonparticipants similar in observable characteristics to IAH beneficiaries but who did not receive home-based primary care. The goals of matching were twofold. First, we sought to minimize nonrandom selection of people in the IAH group by constructing a matched comparison group that appeared similar to the treatment group on key observable characteristics that affect treatment status (such as receipt of home-based primary care from an IAH practice) and outcomes. Subject to that constraint, we then sought to maximize the size of the comparison group to increase statistical efficiency. For the IAH demonstration, key characteristics for matching included those that determined eligibility for the demonstration and measures of health status, health trajectory, and other personal characteristics observable in administrative data that predict health care expenditures. Limiting the comparison group to Medicare beneficiaries who closely matched the observed characteristics of the IAH group may also have reduced

[20] Conducting a preliminary match typically provides a better match in both sites than using a random split, because it ensures that the covariate distribution for the pool of eligible comparison beneficiaries is closely aligned with the covariate distribution for the treatment beneficiaries at each of the two practices.

differences between the two groups on unobserved characteristics if those characteristics were correlated with matching variables.

We conducted matching for the entire IAH group, which consisted of beneficiaries who met the eligibility and attribution criteria based on administrative data. For Year 7, for example, we matched 5,058 IAH beneficiaries on observable characteristics with beneficiaries who were similar and lived in the same geographic area but did not receive home-based primary care. We matched each site separately, including each member of the Richmond-based consortium. We created a comparison group for each practice by estimating a propensity score equation using data for the IAH group and the potential comparison group, and then using the results to find the best matches for each IAH beneficiary.

We used demographic and health-related variables to match beneficiaries in the IAH group with comparison beneficiaries. We used only one measure for exact matching: the number of months since the beneficiary's last hospital admission (one, two or three, or four or more months). Exact matching means that an IAH beneficiary could be matched only to potential comparison beneficiaries who had the same value of that variable. We chose this measure for exact matching because expenditures and use—our key outcomes of interest—tend to be substantially higher in the months after a hospital admission. Preliminary data analyses indicated that adding other exact matching variables would likely result in dissimilarities on other key characteristics, such as disability. Therefore, we chose not to add other exact matching variables. We used two other measures related to eligibility for the demonstration as ordinary matching variables: (1) because a beneficiary can enter the sample at any time in a given year, we used a categorical measure of the month the beneficiary met eligibility criteria (months 1, 2 to 6, or 7 to 12); and (2) because beneficiaries who had an observation stay may have been less acutely ill than those with a hospital admission, we used whether the beneficiary had an observation stay but not a hospital admission in the prior year (Exhibit A.11). We included the following demographic variables in the matching model but did not seek exact matches for them: age (younger than 65, 65 to 79, or 80 or older), gender, race, whether the beneficiary was dually eligible for Medicare and Medicaid, original reason for Medicare eligibility, and number of ADLs (two, three or four, or five or six). We used an indicator variable to identify beneficiaries with missing information for feeding assistance.

We used various measures of health status. We measured individual HCCs using each beneficiary's claims history for the 12 months before the date of eligibility for the demonstration in a given year. Beneficiaries who meet IAH eligibility criteria are at much higher risk of death in a given year than the average Medicare FFS beneficiary, and expenditures in the year before death are sometimes substantially higher than in other years. To increase the likelihood that the comparison group was as similar as

possible to the IAH beneficiaries in health status measures that predict death, we matched the IAH and comparison beneficiaries on risk factors for death. After reviewing the literature on death among Medicare beneficiaries, we selected chronic conditions or diagnoses that were significant predictors of death for use in matching. We included an HCC in the matching equation if Gagne et al. (2011) had identified a diagnosis code as predicting death among elderly Medicare beneficiaries with low income. We collapsed several of the individual HCCs based on the type of condition, frequency in the IAH group, and a relative factor, the last of which represents the contribution of that HCC to the overall HCC risk score.[21,22,23] In Year 7, we added individual HCCs for morbid obesity and immune deficiency, because they were identified as risk factors for negative outcomes resulting from COVID-19 infection (Bosworth et al., 2021). We also used the risk score itself as a matching variable. Additional details about how we calculated the HCC score and HCC indicators are available in Chapter 5 of this appendix.

Exhibit A.11. Variables used in propensity score matching equation

Variable
Eligibility and use
Number of months since most recent hospital admission (1, 2 or 3, 4 or more)
Month of the demonstration year beneficiary met eligibility criteria (1, 2 to 6, 7 to 12)[a]
Whether beneficiary had an observation stay and no hospital admission in prior 12 months
Demographic characteristics
Age: younger than 65, 65 to 79, 80 or older
Gender
Race: White, Black, other, or unknown
Dually eligible for Medicare and Medicaid
Original reason for Medicare entitlement: old age, ESRD or ESRD and disability, disability only
ADLs
Number of ADLs for which beneficiary requires human assistance (2, 3 to 4, 5 to 6)
Whether information about the feeding ADL was missing[b]
Health status
HCC risk score

[21] For example, we combined cirrhosis of the liver (HCC 28) and chronic hepatitis (HCC 29) into a single indicator for matching but did not combine them with end-stage liver disease (HCC 27). Less than 2 percent of the treatment group had cirrhosis of the liver or chronic hepatitis; the relative factor for those conditions was less than half of the relative factor for end-stage liver disease.

[22] Table 9 of the "Announcement of Calendar Year (CY) 2012 Medicare Advantage Capitation Rates and Medicare Advantage and Part D Payment Policies and Final Call Letter" lists the relative factor for each HCC. Available at https://www.cms.gov/Medicare/Health-Plans/MedicareAdvtgSpecRateStats/Downloads/Announcement2012.pdf.

[23] We used software version V2119 to calculate HCC scores for beneficiaries in Year 7, which incorporated version 10 of the *International Classification of Diseases*.

Exhibit A.11 (*continued*)

Variable
Specific HCCs
HCC 2, HCC 6, HCC 46–47, HCC 186, immune deficiency[e]
HCC 8, metastatic cancer and acute leukemia[c]
HCC 9–10, lung and other severe cancers; lymphoma and other cancers
HCC 11–12, colorectal, bladder, and other cancers; breast, prostate, and other cancers and tumors
HCC 18, diabetes with chronic complications
HCC 21, protein-calorie malnutrition
HCC 22, morbid obesity[e]
HCC 27, end-stage liver disease
HCC 28–29, cirrhosis of liver; chronic hepatitis
HCC 46, severe hematological disorders
HCC 48, coagulation defects and other specified hematological disorders
HCC 51, dementia with complications[c]
HCC 52, dementia without complications[c]
HCC 54–55, drug/alcohol psychosis; drug/alcohol dependence
HCC 57–58, schizophrenia; major depressive, bipolar, and paranoid disorders
HCC 70–71, quadriplegia; paraplegia
HCC 72, spinal cord disorders/injuries
HCC 85, congestive heart failure[c]
HCC 96, specified heart arrhythmias
HCC 103–104, hemiplegia/hemiparesis, monoplegia, other paralytic syndromes
HCC 106, atherosclerosis of the extremities with ulceration or gangrene
HCC 107–108, vascular disease with complications; vascular disease
HCC 111, chronic obstructive pulmonary disease
HCC 134, dialysis status[c]
HCC 136–138, chronic kidney disease, stages 3–5[c]
HCC 139–140, chronic kidney disease, stages 1–2 or unspecified; unspecified renal failure
HCC 157–159, pressure ulcer of skin with necrosis through to muscle, tendon, or bone; or with full or partial thickness skin loss
Depression[d]
Anemia
Fluid and electrolyte disorders
Number of chronic conditions (2 to 5, 6 to 9, 10 or more)[d]
Whether beneficiary had a complicating condition or major complicating condition during the most recent hospital admission
Chronically critically ill or medically complex diagnosis

Exhibit A.11 (*continued*)

Note: Exact matching means that an IAH beneficiary can be matched only to a potential comparison beneficiary with the same characteristic. An ordinary matching variable is one used as an independent variable in the matching regression equation.

[a] For pre-demonstration years and Years 1 to 3, month 1 was June or September. For sites that began the demonstration in June 2012, month 1 was June. For sites that began the demonstration in September 2012, month 1 was September. All sites began Years 4 and 5 in October 2015 and October 2016, respectively. Month 1 was January 2019 for Years 6 and 7.

[b] Feeding assessments were not available with home health assessment data at the time of recertification. If the beneficiary had a previous assessment during the study year that was recorded at the time of discharge from home health care, we used the feeding values from that assessment, but sometimes there was no previous discharge assessment.

[c] Gagne et al. (2011) identified these measures as key predictors of death; they are the measures of health status we prioritized most highly when determining which of several alternative matched comparison groups was most appropriate for a particular site in a particular year.

[d] These are chronic condition categories measured by the Chronic Conditions Warehouse.

[e] Added as risk factors for negative outcomes associated with COVID-19 infection (Bosworth et al., 2021)

ADL = activity of daily living; ESRD = end-stage renal disease; HCC = Hierarchical Condition Category.

In addition to the HCCs included in the matching equation based on Gagne et al. (2011) and Bosworth et al. (2021), we included an HCC indicator of pressure ulcers because a large share of the IAH population has poor functional status and could be at higher-than-average risk for a pressure ulcer. We included three other conditions not measured by HCCs: anemia, depression, and electrolyte disorders. Gagne et al. (2011) identified anemia and electrolyte disorders as predictive of death.

We included two other measures of health status using diagnosis codes from the beneficiary's most recent hospital admission in the past year. The first measure indicated whether the Medicare Severity Diagnosis Related Group included a complicating condition or major complicating condition. The second measure indicated whether, according to the diagnosis in the claim, the beneficiary was chronically critically ill or medically complex (Kandilov et al. 2014).

We did not match on previous health care expenditures or health care use, such as type of post-acute care (including source of ADL assessment data) or number of hospital admissions. This was because we were interested in ascertaining whether patients receiving home-based primary care from IAH practices even before the demonstration began were experiencing lower levels of expenditures and use than similarly frail and disabled Medicare beneficiaries who were not receiving in-home primary care. If receiving home-based primary care did result in lower expenditures and hospital use during the pre-demonstration period, matching beneficiaries in the year before the demonstration using utilization experience would yield a comparison group that was healthier, on average, than the IAH group. Thus, a comparison of the two groups would not provide an estimate of the difference in expenditures and use for beneficiaries with equal health care needs resulting from the receipt of home-

based primary care. Rather, it would test whether there were increasing differences in the current year between the level of expenditures and use for an IAH group that would be sicker, on average, than the comparison group with the same level of expenditures and use in the previous year. Similarly, if receiving home-based primary care affected whether a beneficiary received a particular type of post-acute care (such as home health care versus care in a skilled nursing facility), matching beneficiaries on prior use of post-acute care or source of ADL assessment data would be problematic.

We did not use COVID-19 diagnosis as a matching variable in Year 7 because being in the IAH group could have affected the probability that a beneficiary was diagnosed with COVID-19. For example, IAH beneficiaries may disproportionately use home health rather than SNFs for post-acute care, lowering their risk of infection. Or if IAH practices had a more aggressive COVID-19 testing strategy than other primary care providers, then IAH beneficiaries diagnosed with COVID-19 might have had relatively less severe illness than comparison beneficiaries diagnosed with COVID-19. We wanted to be able to observe whether the prevalence of COVID-19 varied for the IAH and comparison groups. As we discuss in Chapter 6 of this appendix, we estimated sensitivity analyses controlling for COVID-19 diagnosis and COVID-19 hospitalization.

4.3 Results of propensity score matching

The standardized difference in means is a standard statistic used to assess similarities between the treatment group and the final matched comparison group (Stuart 2010). The literature suggests that a standardized difference of less than 0.25 is an appropriate threshold for determining that the treatment and comparison groups are well matched on a particular variable (Rubin 2001). We applied a more stringent standard of 0.10 for our matching. We examined the matching results for the variables used in the matching algorithm and additional variables, such as individual HCCs aggregated with other HCCs in the matching equation (for example, cirrhosis of the liver and chronic hepatitis).

Across all 10 sites together (treating the three Richmond-based consortium sites as one), the absolute value of the standardized difference in Year 7 was less than 0.10 on all matching variables and less than 0.10 on all nonmatching variables (Exhibit A.12). All 10 sites individually had standardized differences of less than 0.10 on all of the matching variables. Furthermore, 7 of the sites had standardized differences of less than 0.25 on all of the nonmatching variables. Though COVID diagnosis was not used as a matching variable, the two groups were well-balanced on this indicator; the standardized difference was -0.04.

Exhibit A.12. Characteristics of potential comparison beneficiaries, matched comparison beneficiaries, and IAH beneficiaries, Year 7

Variable	Potential comparison group	Matched comparison group	IAH beneficiaries	Standardized difference
Eligibility for the demonstration				
Number of months since most recent hospital admission, percentage[a]				
One	56.6	37.0	37.0	0.000
Two or three	16.3	19.4	19.4	0.000
Four or more	27.1	43.6	43.6	0.000
Month of the demonstration year that beneficiary met eligibility criteria, percentage[b]				
Month 1	45.4	70.8	69.8	-0.022
Months 2 to 6	27.5	16.5	17.4	0.022
Months 7 to 12	27.1	12.6	12.8	0.005
Observation stay and no hospital admission in previous 12 months	5.3	7.8	7.4	-0.016
Demographic characteristics				
Female	59.4	63.4	63.7	0.007
Dually eligible for Medicare and Medicaid	25.6	41.7	42.5	0.017
Age, percentage				
Younger than 65	11.7	18.1	18.8	0.019
65 to 79	43.0	38.0	38.0	0.000
80 or older	45.3	43.9	43.2	-0.014
Race and ethnicity, percentage				
White	75.1	69.1	69.0	-0.002
Black	17.5	23.6	23.6	0.000
Other	7.5	7.3	7.4	0.003
Original reason for Medicare entitlement, percentage				
Age	73.8	60.2	59.8	-0.009
Disability	23.8	38.4	38.8	0.010
ESRD or ESRD and disability	2.4	1.4	1.4	-0.002
Number of ADLs requiring human assistance, percentage				
Two	10.6	5.7	5.7	0.000
Three or four	29.9	28.0	27.3	-0.015
Five or six	59.5	66.3	67.0	0.014
Missing information about feeding ADL	9.1	18.5	19.8	0.035

Exhibit A.12 (*continued*)

Variable	Potential comparison group	Matched comparison group	IAH beneficiaries	Standardized difference
Health status				
HCC				
HCC risk score	405.0	464.2	469.4	0.026
HCC 2, 6, 46, 47, 186, immune deficiency	29.2	30.5	30.7	0.004
HCC 8, metastatic cancer	5.4	1.8	1.8	0.002
HCC 9–10, lung, lymphoma, and other cancers	6.4	3.9	3.7	-0.008
HCC 11–12, colorectal, bladder, breast, prostate, and other cancers	11.2	8.7	8.5	-0.007
HCC 18, diabetes with chronic complications	39.6	44.6	43.5	-0.021
HCC 21, protein-calorie malnutrition	17.0	26.4	27.8	0.033
HCC 22, morbid obesity	15.7	22.0	21.8	-0.006
HCC 27, end-stage liver disease	2.1	1.6	1.7	0.007
HCC 28–29, cirrhosis of liver and chronic hepatitis	3.3	2.7	2.7	0.001
HCC 46, severe hematological disorders	1.9	1.4	1.3	-0.007
HCC 48, coagulation defects and other specified hematological disorders	19.6	15.5	15.5	0.001
HCC 51, dementia with complications	8.0	13.8	14.1	0.009
HCC 52, dementia without complications	19.8	27.9	27.3	-0.013
HCC 54–55, drug/alcohol psychosis and drug/alcohol dependence	7.9	10.1	10.2	0.002
HCC 57–58, schizophrenia, major depressive, bipolar, and paranoid disorders	25.7	36.6	36.9	0.007
HCC 70–71, quadriplegia, paraplegia	3.2	9.4	10.6	0.046
HCC 72, spinal cord disorders/injuries	3.0	2.5	2.4	-0.009
HCC 85, congestive heart failure	47.8	57.0	56.7	-0.006
HCC 96, specified heart arrhythmias	41.5	37.6	37.2	-0.008
HCC 103–104, hemiplegia/hemiparesis, monoplegia, other paralytic syndromes	11.2	16.6	16.4	-0.004
HCC 106, atherosclerosis of the extremities with ulceration or gangrene	5.7	6.2	6.2	0.002
HCC 107–108, vascular disease with or without complications	46.8	55.8	55.8	-0.001
HCC 111, chronic obstructive pulmonary disease	32.0	40.3	40.2	-0.002
HCC 134, dialysis status	7.3	5.4	5.3	-0.004

Exhibit A.12 (*continued*)

Variable	Potential comparison group	Matched comparison group	IAH beneficiaries	Standardized difference
HCC 136–138, chronic kidney disease, stages 3–5	9.5	13.4	13.1	-0.010
HCC 139–140, chronic kidney disease stages 1–2, unspecified renal failure	4.2	6.4	6.5	0.004
HCC 157–159, pressure ulcer of skin with necrosis or skin loss	10.2	21.1	22.9	0.047
Number of chronic conditions measured by Chronic Conditions Warehouse				
Fewer than six	13.5	10.8	11.2	0.011
Six to nine	49.1	44.4	44.4	-0.001
More than nine	37.4	44.7	44.4	-0.006
Other measures of health status				
Anemia[c]	18.6	20.0	19.8	-0.006
Depression	48.7	55.8	55.8	0.001
Fluid and electrolyte disorders[c]	43.5	44.6	44.5	-0.001
Diagnosis of chronically critically ill or medically complex[d]	41.2	43.7	44.1	0.008
Complicating condition or major complicating condition during the most recent hospital admission	64.2	63.0	63.3	0.007

Source: Medicare claims and enrollment data obtained from the Chronic Conditions Warehouse for IAH and matched comparison group beneficiaries in all IAH practices that participated in demonstration Year 7.

Notes: The final sample sizes in Year 7 were 5,058 IAH beneficiaries and 20,831 matched comparison beneficiaries. The number of weighted matched comparison beneficiaries equaled the number of IAH beneficiaries.

[a] Variable used for exact matching.

[b] Month refers to the first month in the demonstration year after the beneficiary met eligibility criteria. For example, if a beneficiary had a qualifying admission and rehabilitation services in one or more months before the demonstration, the month 1 group included that person. For all sites in Year 7, month 1 was January.

[c] Measured using claims from the most recent hospital admission or observation stay in the year before the demonstration eligibility date. We drew diagnosis codes for these conditions from Gagne et al. (2011).

[d] Measured using diagnoses from the most recent hospital admission in the year before the demonstration eligibility date. We drew diagnosis codes from Kandilov et al. (2014).

ADLs = activities of daily living; ESRD = end-stage renal disease; HCC = Hierarchical Condition Category.

All subgroups for which we estimated effects of IAH were well-balanced. We examined balance for the following subgroups: number of ADLs requiring human assistance (2 to 4 and 5 to 6), race (Black, White, and other), dual status (dual and non-dual), age (under 85 and ages 85 or older), chronic condition count (9 or fewer and more than 9), and original reason for Medicare entitlement (age and reason other than age). Among these 13 subgroups, 10 had zero matching variables with standardized differences that were greater than 0.10 or less than -0.10. No subgroup had a matching variable with a standardized difference that exceeded 0.15.

4.4. Number of beneficiaries and eligible months

Beneficiaries in both the IAH group and comparison group were analyzed from the month they became eligible for the demonstration and observed for the remaining months in a given demonstration year. Over the seven years of the demonstration, the number of IAH beneficiaries varied; for each IAH beneficiary, we matched up to five comparison beneficiaries. On average, each IAH beneficiary matched to four comparison beneficiaries.

Across the demonstration years, the average number of eligible months for the comparison beneficiaries was slightly smaller than among the IAH beneficiaries (Exhibits A.13 through A.15). This difference arose because the comparison beneficiaries were more likely to die during the demonstration year than the IAH beneficiaries (the rate of death in Year 7 was 19.4 percent for IAH beneficiaries and 27.0 percent for comparison beneficiaries), and the IAH beneficiaries were more likely to qualify for the demonstration earlier in the 12-month period than the comparison beneficiaries. To address any possible concerns that this difference might cause, we incorporated an eligibility fraction into the weighting design for regressions, in which the eligibility weight reflected the number of months eligible for the demonstration in a given year. For example, a beneficiary eligible for the demonstration for 6 months in Year 7 had half the weight of a beneficiary eligible for the demonstration for 12 months in Year 7. Using an eligibility fraction in the weight ensured that each beneficiary's contribution to the estimation was proportionate to how long we observed that person during a given year. In addition, we used the following as matching and control variables: number of months since most recent hospital admission and month of the demonstration year that the beneficiary met the eligibility criteria. In this way, we accounted for differences in the time between when beneficiaries met the service use criteria required for demonstration eligibility and their eligibility date. Those who qualified in the first month may have met both the service use criteria up to one year before the demonstration year began, and those who qualified in later months met at least one of the two service use criteria in the month immediately before the eligibility date. Chapter 6 of this appendix provides additional details about weights and control variables.

Exhibit A.13. Analysis sample, by years, practices that participated in Years 1 to 5

	Two years before the demonstration	One year before the demonstration	Year 1	Year 2	Year 3	Year 4	Year 5
Number of IAH beneficiaries	6,837	7,367	8,216	7,266	7,564	9,504	9,958
Total number of eligible months for IAH beneficiaries	65,781	70,591	79,396	69,768	72,215	90,223	95,003
Average number of eligible months per IAH beneficiary	9.6	9.6	9.7	9.6	9.5	9.5	9.5
Number of comparison beneficiaries	29,517	31,888	33,916	32,248	31,259	38,365	41,387
Total number of eligible months for comparison beneficiaries	264,558	286,314	303,770	293,081	278,015	335,250	363,251
Average number of eligible months per comparison beneficiary	9.0	9.0	9.0	9.1	8.9	8.7	8.8

Source: Medicare claims and enrollment data obtained from the Chronic Conditions Warehouse for IAH and matched comparison group beneficiaries in all IAH practices that participated in Year 5.

Note: This table reflects the sample we used to estimate the effects of the IAH payment incentive in Years 1 to 5.

Exhibit A.14. Analysis sample, by years, practices that participated in Year 6

	Two years before the demonstration	One year before the demonstration	Year 1	Year 2	Year 3	Year 4	Year 5	Year 6
Number of IAH beneficiaries	5,524	5,826	7,037	6,266	6,626	8,487	9,006	7,993
Total number of eligible months for IAH beneficiaries	53,071	55,802	67,596	59,993	63,068	80,595	85,611	77,224
Average number of eligible months per IAH beneficiary	9.6	9.6	9.6	9.6	9.5	9.5	9.5	9.7
Number of comparison beneficiaries	23,839	25,089	28,325	27,315	26,630	33,348	36,707	32,496
Total number of eligible months for comparison beneficiaries	212,510	225,531	251,398	247,050	235,762	291,711	320,428	291,352
Average number of eligible months per comparison beneficiary	8.9	9.0	8.9	9.0	8.9	8.7	8.7	9.0

Source: Medicare claims and enrollment data obtained from the Chronic Conditions Warehouse for IAH and matched comparison group beneficiaries in all IAH practices that participated in Year 6.

Note: This table reflects the sample we used to estimate the effects of the IAH payment incentive in Year 6.

Exhibit A.15. Analysis sample, by years, practices that participated in Year 7

	Two years before the demonstration	One year before the demonstration	Year 1	Year 2	Year 3	Year 4	Year 5	Year 6	Year 7
Number of IAH beneficiaries	4,693	4,944	6,026	5,034	5,202	6,633	6,896	5,907	5,058
Total number of eligible months for IAH beneficiaries	45,364	47,635	58,387	48,555	49,830	63,668	66,346	56,890	49,301
Average number of eligible months per IAH beneficiary	9.6	9.6	9.7	9.6	9.6	9.6	9.6	9.6	9.7
Number of comparison beneficiaries	20,395	21,536	24,240	22,373	20,800	26,026	28,489	24,526	20,831
Total number of eligible months for comparison beneficiaries	182,421	194,836	216,825	204,016	185,186	228,864	251,438	218,136	182,112
Average number of eligible months per comparison beneficiary	8.9	9.0	8.9	9.1	8.9	8.8	8.8	8.9	8.7

Source: Medicare claims and enrollment data obtained from the Chronic Conditions Warehouse for IAH and matched comparison group beneficiaries in all IAH practices that participated in Year 7.

Note: This table reflects the sample we used to estimate the effects of IAH in Year 7.

5. Medicare data and measures

In this chapter, we describe the data sources and measures we used in our analyses of the effect of the demonstration.

We constructed our yearly analytic files with observations at the beneficiary year level. We drew data for determining demonstration eligibility and measuring outcomes in the analytic files from several sources (Exhibit A.16). We accessed all data through the Chronic Conditions Warehouse Data Enclave.

Exhibit A.16. Data sources

Data	Demographic characteristics	Chronic conditions	Activities of daily living	Service use: Demonstration eligibility	Service use: Outcome measures	Health outcomes
Medicare enrollment database	✓					✓
Master beneficiary summary file		✓				
Inpatient claims				✓	✓	
Outpatient claims				✓	✓	
Clinician/supplier claims				✓	✓	
Home health agency claims				✓	✓	
Skilled nursing facility claims				✓	✓	
Hospice claims				✓	✓	
Durable medical equipment claims					✓	
IRF-PAI[a]			✓			
MDS			✓		✓	
OASIS			✓			
Timeline file						✓

Source: Medicare claims and enrollment data.

[a] This instrument includes inpatient rehabilitation hospitals and rehabilitation units and excludes long-term care hospitals.

IRF-PAI = Inpatient Rehabilitation Facility Patient Assessment Instrument; MDS = Minimum Data Set; OASIS = Outcome and Assessment Information Set.

5.1. HCC score and indicators

To account for differences in health status and the differential risks of incurring high Medicare expenditures, we used the CMS-HCC risk-adjustment model to create HCC scores and indicators. To estimate the HCC scores, we used a 12-month look-back period for Medicare claims to obtain diagnosis information. Because the claims-based eligibility dates for IAH and comparison beneficiaries can vary for a specific pre-demonstration or demonstration year, the 12-month look-back period also varied depending on the beneficiaries' eligibility dates. For each beneficiary in the IAH and comparison groups, we estimated the HCC score by using the publicly available HCC software (CMS 2019) and information on demographics, Medicare eligibility, and dual eligibility status, as well as Medicare claims for the 12 months before the person's claims-based eligibility date. We used fewer than 12 months of Medicare claims if a beneficiary was not enrolled in Medicare for all 12 months.

As with all previous years of the demonstration, Year 7 HCC scores were calculated using HCC Version 21. This model was developed and calibrated for the Program of All-Inclusive Care for the Elderly population, which resembles the IAH-eligible population in being sicker and frailer than the average Medicare beneficiary. Though CMS ended support for Version 21 in 2019, we chose to continue using it in Year 7 rather than switching to the updated Version 22 to maintain consistency across years. HCC Version 22 uses a different model with different HCC categories and different model coefficients, so data generated using Version 22 may not be comparable to data generated using Version 21. Maintaining consistency is important, because our difference-in-differences approach uses data from all previous demonstration and pre-demonstration years when estimating the impacts in the latest demonstration year. It is possible that changing from Version 21 to Version 22 could affect IAH and comparison beneficiaries differently, which would pose a risk of bias in the study results. Although CMS ended support for Version 21, it continues to maintain the Version 21 code sets to include new codes to ensure they are up-to-date.

CMS has separate HCC models for beneficiaries residing in the community and those residing in an institution. We used the HCC score estimated by the community model for all beneficiaries in our sample. Beneficiaries cannot reside in an institution when they become eligible for the demonstration, so we did not use scores generated by the institutional model for any beneficiary. We also did not use the demographics-only model for new enrollees. Because of the service use requirements for the demonstration, all IAH-eligible beneficiaries had some claims history during the previous 12 months. Using any available diagnosis information in the HCC model should have provided a score that captured health status better than a demographics-only model. The specific scale of the HCC score should not have affected propensity score matching if the score was estimated similarly for both IAH

and potential comparison beneficiaries; thus, we did not normalize or rescale HCC scores.

5.2. Dual eligibility

When we did propensity score matching for the full sample in all demonstration and pre-demonstration years, we measured dual eligibility using the monthly Part A and Part B state buy-in variables on Medicare enrollment data. We used this approach because Medicaid enrollment data were not available promptly enough for us to define dual eligibility using those data, and the evaluation began several years before the Medicare Beneficiary Summary File included monthly Medicare-Medicaid dual eligibility codes. If a beneficiary had state buy-in for Part A, Part B, or both in any month in a pre-demonstration or demonstration year, we identified that person as being dually eligible in that year. We used the same measure of dual eligibility as a control variable in the regression models for Medicare expenditures and other Medicare claims-based outcomes.

5.3. Outcome variables

We measured outcomes for the period that beneficiaries were eligible during a given pre-demonstration or demonstration year, which started from the date of eligibility through the end of the demonstration year or date of death. Therefore, the eligibility period differed across beneficiaries, depending on their eligibility start dates and death dates.

We used four groups of measures for the regression analysis of outcomes in the demonstration based on Medicare Part A and Part B claims as well as the Medicare enrollment database: (1) Medicare expenditures, (2) hospital care use, (3) quality of care, and (4) health outcomes (Exhibit A.17). We measured these outcomes for the number of months a beneficiary was observed in a study year, starting with the first day of the first month after the beneficiary met all eligibility criteria in each year based on our analysis of Medicare enrollment and administrative data.

We measured all claims-based outcomes at the beneficiary level in that particular study year. For expenditures, we measured each outcome PBPM. For example, if a beneficiary was alive and in Medicare FFS for four months from the demonstration eligibility date through the end of the year, we divided expenditures during those four months by four to measure expenditures PBPM. We annualized claims-based outcomes other than expenditures and binary measures (such as the likelihood of unplanned readmission or death). For example, if beneficiaries had four hospital admissions and an eligibility weight of 0.5 (because they were eligible for the demonstration for 6 of 12 months in the demonstration year), the annualized number of hospital admissions would be eight.

Exhibit A.17. Measures of Medicare expenditures, hospital use, quality of care, and health outcomes used in regressions

Measure
Medicare expenditures per beneficiary per month
Total
Inpatient
Home health service[a]
Outpatient
Skilled nursing facility
Clinician/supplier
Hospice
Durable medical equipment
Hospital use
Number of hospital admissions per beneficiary per year[b]
Number of hospital admissions preceded by an ED visit per beneficiary per year
Number of ED visits per beneficiary per year
Quality of care
Number of potentially avoidable hospital admissions per beneficiary per year (AHRQ PQI)[b]
Number of potentially avoidable outpatient ED visits per beneficiary per year (AHRQ PQI)[c]
Probability (as a percentage) of having a qualifying hospital discharge and unplanned readmission within 30 days of discharge[d]
Health outcomes
Death within the study year
Entry into institutional long-term care within the study year

Notes: We constructed measures using data from the date the beneficiary became eligible in the demonstration year through the end of that demonstration year. Following the CMMI Priority Measures for Monitoring and Evaluation, we did not truncate expenditure measures. Instead, we risk-adjusted, annualized, and weighted them to reflect partial year observations. We did not price standardize the expenditure measures.

[a] Total home health expenditures include all care provided under the home health benefit. Claims for therapy appear only in the outpatient file.

[b] This includes hospital admissions and observation stays.

[c] We measured this as specified in the CMMI Priority Measures for Monitoring and Evaluation.

[d] Qualifying hospital discharges include discharges for patients who were enrolled in Medicare FFS, discharged from nonfederal acute care hospitals, alive at the time of discharge, and not transferred to another acute care facility. Home-based primary care and the demonstration may affect whether a beneficiary has a qualifying hospital discharge in a particular year. Such an effect could lead to estimating biased rates of readmission for the IAH and comparison groups if readmission is defined only for beneficiaries who had a qualifying hospital discharge, as recommended by the CMMI Priority Measures for Monitoring and Evaluation. Thus, we defined the readmission measure using all beneficiaries in the denominator rather than limiting it to beneficiaries with a qualifying discharge. For example, if home-based primary care or the demonstration reduces the likelihood of having a qualifying hospital discharge, IAH beneficiaries who have such a discharge may be sicker on average than comparison beneficiaries who have such a discharge. Being sicker could lead to an increased risk of readmission.

Exhibit A.17 (*continued*)

AHRQ = Agency for Healthcare Research and Quality; CMMI = Center for Medicare & Medicaid Innovation; ED = emergency department; FFS = fee-for-service; PQI = Prevention Quality Indicator.

Potentially avoidable hospital admissions and outpatient ED visits. Potentially avoidable hospital use occurs when ambulatory care may have prevented or reduced the need for a hospital admission or ED visit. We measured a beneficiary as having a potentially avoidable hospital admission or ED visit if the principal diagnosis for the hospital admission or ED visit was an ambulatory care-sensitive condition. We based our definition of ambulatory care-sensitive conditions on the Agency for Healthcare Research and Quality's Prevention Quality Indicator 90, which includes the following conditions for 2020: diabetes short-term complications, diabetes long-term complications, uncontrolled diabetes, lower-extremity amputation among diabetics, chronic obstructive pulmonary disease or asthma in older adults, hypertension, heart failure, community-acquired pneumonia, and urinary tract infection. The measure of potentially avoidable ED visits included only outpatient ED visits (that is, ED visits not accompanied by an admission). We excluded ED visits that led to a hospital admission because the principal diagnosis on the inpatient claim would not necessarily be the ambulatory care-sensitive condition leading to the ED visit.[24]

Hospital admission preceded by an ED visit. In addition to measuring total and potentially avoidable hospital admissions, we measured hospital admissions preceded by an ED visit. The IAH demonstration could have different effects on the various hospital admission measures. Also, hospital admissions preceded by an ED visit are likely to be a larger component of total hospital admissions for beneficiaries who meet IAH eligibility requirements than for the general population of Medicare FFS beneficiaries.

Outpatient ED visits. Our measure of emergency care was the number of outpatient ED visits (including those visits that led to an observation stay). This measure included cases in which a beneficiary was transferred to a different hospital for admission and may have included some cases in which a hospital billed ED and inpatient services separately.

Unplanned readmission within 30 days of discharge. The unplanned readmission measure indicated whether the beneficiary had at least one unplanned readmission within 30 days of a qualifying hospital discharge. Qualifying hospital discharges for

[24] ED visits appear in Medicare inpatient and outpatient claims. A beneficiary whose ED visit led to a hospital admission would not have a separate claim in the outpatient file; the ED claim would be part of the hospital claim in the inpatient file, which would have diagnoses that reflect the hospital stay. Using inpatient claims to measure potentially avoidable ED visits that led to hospital admission poses two problems. First, the diagnosis that led someone to the ED may be different from the diagnoses on the inpatient claim (for example, a beneficiary visits the ED because of shortness of breath but is later admitted because of another underlying factor). Second, hospital admissions with potentially avoidable diagnoses are counted in the potentially avoidable hospital admission measure. If we also counted them as potentially avoidable ED visits, we would double count the utilization.

the readmission measure included discharges from nonfederal acute care hospitals for patients who were enrolled in Medicare FFS, alive at the time of discharge, and not transferred to another acute care facility. The qualifying discharges included patients discharged to nonacute care settings. Index discharges did not include admissions to Prospective Payment System–exempt cancer hospitals or admissions for patients without at least 30 days of post-discharge enrollment in FFS Medicare Parts A and B (unless a patient was enrolled in FFS but died within 30 days), patients discharged against medical advice, primary psychiatric diagnoses, rehabilitation, and medical treatment of cancer.

We excluded planned readmissions from this measure. To identify them, we followed the approach used by CMS's hospital-level 30-day risk-standardized readmission measure developed by the Yale New Haven Health Services Corporation/Center for Outcomes Research & Evaluation (2020). Unlike the Yale measure, our list of procedure codes to identify planned readmissions did not include codes that apply only to all-payer populations.

All beneficiaries who had a qualifying hospital discharge and an unplanned readmission within 30 days were identified as having an unplanned readmission. Therefore, the measure provided an estimate of the combined effect of the demonstration on whether patients had a qualifying hospital discharge and, if so, whether they had an unplanned readmission within 30 days.

Entry into institutional long-term care. The institutional long-term care measure identified beneficiaries who had at least one episode of long-term care spanning 90 or more days during a given study year. An episode of long-term care began when a beneficiary entered a skilled or unskilled nursing facility and ended when the beneficiary spent more than 14 consecutive days in the community or the study year ended. We evaluated each study year (demonstration or pre-demonstration year) separately. In other words, we required a beneficiary to have a 90-day episode of institutional long-term care during a single study year to identify that person as entering long-term care in that study year. Beneficiaries in long-term care for the entire time they were otherwise eligible for the demonstration in a given year could not be in the IAH group for the evaluation in that year.

We created this measure using the Timeline file, which combines data from claims (inpatient, SNF, and home health) and assessment data (MDS and OASIS) to flag a beneficiary's residency status for each day of a calendar year. The daily residency flag can contain one of the following values: I (inpatient), S (SNF), M (MDS, which includes nursing home days not paid by Medicare), C (community, which includes days identified by home health claims or OASIS assessment data and days with no claims or assessment data), D (dead), or blank (not Medicare eligible). We considered all days flagged with a C, H, or O, or that were blank, to be community days. The small amount of missing OASIS data and decrease in number of MDS assessments that we

discussed in Chapter 3 of this appendix pose little risk of error for the measure of entry into institutional long-term care. This is because (1) OASIS assessments are used to identify community days in the Timeline file, and community is residency status assigned to a beneficiary when data are missing; and (2) CMS has reported that most facilities submitted the MDS within the usual timeframe despite the timeframe waiver available during the COVID-19 public health emergency in 2020.

Potential episodes of long-term care began when the beneficiary entered a long-term care institution (a daily status of S or M) in a given study year. Inpatient days that occurred during a potential episode of long-term care were considered part of the long-term care episode. An episode of long-term care, however, could not begin with an inpatient stay. For example, an inpatient day that immediately preceded the beneficiary's first SNF or MDS day did not count toward an episode of long-term care, but an inpatient day that occurred the day after a SNF day did. We counted days in the community that occurred during an episode of long-term care toward the 90-day requirement as long as there were no more than 14 consecutive community days and the beneficiary reentered an institution—a daily status of S, M, or I—on or before a 15th community day.

5.4. Measures of visits and home health use

We used clinician/supplier claims from Year 7 to measure visits along three dimensions:

1. Mode of visit: in-person, telehealth, and telephone. A telehealth visit requires real-time audio and video communication between the clinician and the patient whereas a telephone visit requires only real-time audio. We identified telephone visits using procedure codes. Unlike telephone visits, telehealth visits use the same procedure code that would have been used had the visit occurred in person. Therefore, we identified telehealth visits based on the telehealth modifier.
2. Location of visits: home and office/clinic. Home visits include visits that took place in private homes, assisted living facilities, group homes, and custodial care facilities. Unless otherwise specified, home and office visits include in-person visits and telehealth visits. Because we did not use outpatient claims to measure visits to federally qualified health centers and rural health clinics, the measure of office/clinic visits might slightly undercount visits. This limitation likely applies to the comparison group more than to the IAH group, because the IAH group received care from IAH practices that submit clinician/supplier claims.
3. Type of clinician: primary care (including primary care physicians, physician assistants [PAs], and NPs and other advanced practice nurses) and specialty physicians. We used provider specialty codes from claims to identify clinician type. There are two limitations to these measures. First, because we could not identify the specialty for PAs and NPs, the measure of primary care visits might

slightly overcount primary care, and the measure of specialty care visits might slightly undercount specialty care. This limitation likely applies to the comparison group more than to the IAH group, because most visits we identified from NPs and PAs for IAH beneficiaries were provided by IAH practices, all of which provide primary care exclusively. Second, as noted previously, we did not use outpatient claims to measure visits to federally qualified health centers and rural health clinics. Therefore, the measure of primary care visits might slightly undercount primary care visits—a limitation that likely applies to the comparison group more than to the IAH group.

In all cases, we used procedure codes to limit visits to evaluation and management services, transitional care management services, annual wellness visits, advance care planning, and cognitive assessment and care plan services. We limited visits to those made in an outpatient clinic, office, or the beneficiary's home because we wanted to focus on visits that occurred in the community. To limit to these settings, we excluded claims with place of service codes indicating that the service was provided in an inpatient hospital, emergency department, ambulatory surgical center, birthing center, nursing facility or skilled nursing facility, ambulance, psychiatric facility, inpatient rehabilitation facility, or treatment facility for end-stage renal disease.

Unless otherwise stated, results in Chapter 2 of the report for primary care visits and primary care spending reflect all modes of care (in-person, telehealth, and telephone) and locations (home and office/clinic) for visits provided by primary care physicians, PAs, and NPs. Likewise, specialist visits and specialist spending reflect all modes of care (in-person, telehealth, and telephone) and locations (home and office) for visits provided by specialty physicians.

As described in Chapter 2 of the report, we used home health claims from Year 7 to measure the share of beneficiaries who used any home health and spending for home health services.

6. Estimation of demonstration impacts

We used a difference-in-differences model to estimate the impact of IAH in each demonstration year. Our difference-in-differences impact estimates measured the difference in a given outcome between the year before the demonstration started and any demonstration year for beneficiaries comprising the IAH group relative to the difference during the same period for beneficiaries comprising the comparison group. We implemented the difference-in-differences model using two approaches: a frequentist approach and a Bayesian approach.

In this chapter, we describe changes in the sample starting in Year 7 and how those differences affect our estimation, and then we discuss how the interpretation of estimated effects in Year 7 differs from Years 1 to 6. We then present all methodology related to the frequentist analyses, followed by describing aspects of the Bayesian analysis that differed from the frequentist analyses.

6.1. Samples used to estimate effects in each demonstration year

Our primary estimates of the effect of IAH on outcomes are based on three samples of IAH practices from separate regressions, depending on the years being estimated. For Years 1 to 5, we estimated effects using a sample containing the 14 sites that participated through Year 5 of the demonstration; for Year 6, we estimated effects using a separate regression from a sample containing the 12 sites that participated in Year 6. Similarly, for Year 7, we estimated effects using a separate regression from a sample containing the 10 sites that participated in Year 7. (See Chapter 3 of this appendix for additional information about why we did not retain IAH practices after they stopped participating in the demonstration.) In addition to the primary impact estimate in Year 7, we also use results from the regression containing only the 10 sites that participated in Year 7 to examine whether the estimated effect of the demonstration changed from Years 6 to 7 for sites that participated in both years.

6.2. Interpretation of estimated effects

In prior reports, including Year 6, we interpreted the estimated effects from our difference-in-differences model as estimated effects of the IAH payment incentive. As we note in several places throughout this report, Year 7 effects from the same models should be interpreted differently—specifically, as effects of IAH *during the first year of the COVID-19 pandemic*. When examining effects for Years 1 to 6, we assumed that the relative effectiveness of care for IAH and comparison beneficiaries remained constant from before the demonstration through Year 6 (2019). This assumption means that any estimated effect of IAH in Years 1 to 6 would have been due to changes made by IAH practices because of the demonstration payment incentive. In Year 7, the first year of the COVID-19 pandemic (2020), the Year 7 effect estimate

includes any changes to the relative effectiveness of care for IAH and comparison beneficiaries if any of the following occurred:

- IAH practices may have changed care delivery.
- Care delivery approaches that IAH practices have used since before the demonstration began may have been more effective.
- The comparison group may have experienced changes in care that did not affect the IAH group because of efforts by IAH practices.

In addition to potential changes in the relative effectiveness of care, unmeasured factors might have played a larger role in Year 7 than in prior demonstration years. The difference-in-differences methodology removes the effect of time-invariant unmeasured factors that influence outcomes. However, if those factors or their influence on outcomes changed over time (in either direction), then the estimated effect may be different from what it would have been if we had been able to accurately measure and control for those unmeasured factors as part selecting the comparison group and estimating regressions.

In summary, results from Year 7 should be interpreted as estimated effects of IAH during the first year of the COVID-19 pandemic. In addition, there could be an increased risk of confounding in the estimated effects in Year 7 due to changes in unmeasured factors; even if any such changes were small individually, they could have accumulated to account for a non-trivial share of the estimated effects of IAH in Year 7. In Chapter 4 of the report, we discuss interpretation of estimated effects in more detail.

6.3. Frequentist difference-in-differences model

6.3.1. Model specification for continuous and count outcomes

We estimated the impacts of the demonstration by comparing the regression-adjusted differences in outcomes between the IAH treatment and comparison groups in the pre- and post-demonstration periods. We used a difference-in-differences estimation strategy to test for differential changes in all claims-based outcomes between the IAH and comparison groups during the two pre-demonstration years and the first seven years of the demonstration. Equation (1) shows the model we estimated for each outcome in Year 7:

(1) $$Y_{it} = \alpha + X_{it}\beta + \tau \bullet treatment_{it} + \gamma_{-1}\, PD_1 + \gamma_1\, DY_1 + \gamma_2\, DY_2 + \gamma_3\, DY_3 + \gamma_4\, DY_4 + \gamma_5\, DY_5 + \gamma_6\, DY_6 + \gamma_7\, DY_7 + \theta_{-1} treatment_{it} \bullet PD_1 + \theta_1 treatment_{it} \bullet DY_1 + \theta_2 treatment_{it} \bullet DY_2 + \theta_3 treatment_{it} \bullet DY_3 + \theta_4 treatment_{it} \bullet DY_4 + \theta_5 treatment_{it} \bullet DY_5 + \theta_6 treatment_{it} \bullet DY_6 + \theta_7 treatment_{it} \bullet DY_7 + \varepsilon_{it}$$

Where Y_{it} is the claims-based outcome measured for a beneficiary *i* in pre-demonstration or demonstration year *t*; α is a constant term; X_{it} is a set of

beneficiary characteristics measured in the index year; PD_1 is an indicator for pre-demonstration Year 1 (that is, two years before the start of the demonstration, with the year immediately preceding the demonstration serving as the reference or omitted category); $DY_1 - DY_7$ are a set of indicators for each post-demonstration year; $treatment_{it}$ is an indicator variable for being in an IAH practice; and ε_{it} is a random error term. As we describe later in this chapter, the set of beneficiary characteristics included in X_{it} were largely the same as the variables used for matching; they controlled for any remaining differences between the IAH and matched comparison groups in these characteristics.

The key parameters are $\theta_1 - \theta_7$, which constitute the difference-in-differences coefficients; they are the change in an outcome from the year before the demonstration to each year after the demonstration for the IAH group, net of the change in outcome for the comparison group during the same period. Separate estimates for each year (that is, one θ per year) allowed for nonlinearities in such trends. Last, the parameter θ_{-1} captures the differential change in outcome between the IAH and matched comparison groups during the two pre-intervention years. We use θ_{-1} to examine whether the two groups were on the same outcome trajectories before the demonstration; we discuss this test later in this chapter.

In cases in which we estimated a linear model, such as total Medicare expenditures, the difference-in-differences coefficients $\theta_1 - \theta_7$ equaled the difference-in-differences impact estimates. In cases in which we used non-linear models, such as a negative binomial regression for the number of hospital admissions, we transformed $\theta_1 - \theta_7$ into difference-in-differences impact estimates using the following steps using the estimated impact in Year 7 as an example:

1. Using the coefficients obtained from Equation (1), we calculated the average outcomes for the IAH treatment and comparison groups in each year. We adjusted the yearly average outcomes for both groups to reflect the covariate distribution of the IAH group in the latest demonstration year (Year 7). For example, we used the mean covariate values of the Year 7 IAH group to generate two estimates of predicted total Medicare expenditures in the year before the demonstration: one estimate assumed that beneficiaries received home-based primary care in that year (the IAH treatment group estimate), and one assumed that beneficiaries did not receive home-based primary care in that year (the comparison group estimate).
2. We calculated the difference of the regression-adjusted outcome for the IAH group and matched comparison group in Year 7.
3. We calculated the change in the difference between the IAH and matched comparison groups in Year 7 relative to the difference in the year before the

demonstration. We refer to this estimate as the difference-in-differences impact estimate.

Our difference-in-differences impact estimates measured the change between two differences: the pre- and post-demonstration difference for IAH beneficiaries, and the pre- and post-demonstration difference for comparison beneficiaries. This method isolated the impact of the demonstration by accounting for two factors that affected outcomes. First, it accounted for the difference in outcomes between IAH and comparison beneficiaries before the demonstration, controlling for differences in observed beneficiary characteristics. Second, it accounted for changes in outcomes during the demonstration caused by factors unrelated to the demonstration that affected IAH and comparison beneficiaries over time. However, as we discussed previously in this chapter, the interpretation of the impact estimates differed during Year 7—the first year of the COVID-19 pandemic—compared with Years 1 to 6.

In addition to reporting all difference-in-differences estimates in absolute terms, we calculated the impacts in percentage terms by dividing the impact estimate for an outcome by the unadjusted IAH group mean for that same outcome in the year before the demonstration. The percentage impact helped us to interpret whether the magnitude of an impact in a given year is meaningful in practical terms.

We used linear regressions for expenditures. We used negative binomial regressions for the number of hospital admissions and ED visits to account for over-dispersion of counts, and zero-inflated negative binomial regressions for the number of potentially avoidable hospital admissions and outpatient ED visits to account for both over-dispersion and the large percentage of beneficiaries with no admissions or outpatient ED visits during the time period.

For all outcomes, we adjusted standard errors for clustering at the practice level for the IAH group and at the beneficiary level for the comparison group (which we refer to below as hybrid clustering). We estimated the effect of the demonstration on all outcomes using two weighting schemes, which we refer to as beneficiary weighting and practice weighting. Sections 6.3.4 and 6.3.5 of this appendix describe clustering and weighting in detail.

6.3.2. Model specification for death

We used survival modeling techniques to estimate whether the demonstration had an effect on the probability of a beneficiary dying within the demonstration year. The advantage of this approach relative to a logistic regression model is that it allowed us to use a flexible functional form to account for some beneficiaries becoming eligible after the beginning of the demonstration year and thus having shorter periods of observation relative to other beneficiaries. We used the accelerated failure time hazard specification to estimate a survival-time model in Equation (2) as follows:

(2)

$$\log(T_{it}) = X_{it}\beta + \tau \bullet treatment_{it} + \gamma_{-1}\, PD_1 + \gamma_1\, DY_1 + \gamma_2\, DY_2 + \gamma_3\, DY_3 + \gamma_4\, DY_4 + \gamma_5\, DY_5 + \gamma_6\, DY_6 + \gamma_7\, DY_7 + \theta_{-1} treatment_{it} \bullet PD_1 + \theta_1 treatment_{it} \bullet DY_1 + \theta_2 treatment_{it} \bullet DY_2 + \theta_3 treatment_{it} \bullet DY_3 + \theta_4 treatment_{it} \bullet DY_4 + \theta_5 treatment_{it} \bullet DY_5 + \theta_6 treatment_{it} \bullet DY_6 + \theta_7 treatment_{it} \bullet DY_7 + z_{it}$$

Where T_{it} denotes the number of days that beneficiary *i* survived in demonstration year *t* subsequent to that individual's eligibility date in that year; X_{it} includes the same set of beneficiary characteristics measured in the index year as in Equation (1); PD_1 is an indicator for two years before the demonstration; $DY_1 - DY_7$ are a set of indicators for each post-demonstration year; and $treatment_{it}$ is an indicator variable for being in an IAH practice. The term z_{it} is an error term with a distribution $f(\cdot)$.

The model in Equation (2) accounts for the exact survival time not being observed for beneficiaries who did not die at the end of a given demonstration year (that is, right censoring) and the survival time not being measured from the beginning of the demonstration year for beneficiaries who entered the study sample late (that is, left truncation). We estimated the model using the maximum likelihood method with a generalized gamma distribution for $f(\cdot)$ to allow for the possibility of non-monotonic hazard functions.[25] We used matching weights to account for the number of matched comparisons per IAH beneficiary so that the two groups were the same size. We adjusted standard errors using the hybrid clustering approach, which we describe in detail below.

After estimating the survival regression, we transformed $\theta_1 - \theta_7$ into difference-in-differences effect estimates, following steps similar to those we used for estimating impacts for other outcomes. Specifically, we obtained the regression-adjusted average death rate (that is, one minus the probability of survival by the end of the demonstration year) for the IAH and comparison groups in each year. We adjusted the yearly average outcomes for both groups to reflect the covariate distribution of the IAH group in Year 7. Finally, we estimated the difference-in-differences impact by calculating the difference of the regression-adjusted death rate for the IAH group and matched comparison groups in that year relative to the difference between the two groups in the year before the demonstration.

[25] To inform our choice of the survival function, we compared the goodness of fit of models using different distributions. We considered five types of parametric survival distributions: (1) Weibull, (2) log logistic, (3) log normal, (4) generalized gamma, and (5) Gompertz. In choosing the final model, we analyzed the log likelihood, the Akaike information criterion, and the Bayesian information criterion across these different models, selecting the distribution that consistently produced the smallest value on these metrics.

As part of the outputs from the survival regression, we obtained the predicted death rate for each beneficiary during a given demonstration year based on the individual's treatment status and baseline characteristics. This predicted death rate, denoted as h_{it}, then fed into the estimation model for other binary outcomes, which we describe next.

6.3.3. Model specification for other binary outcomes

In addition to death, we estimated the impact of the demonstration on two other binary outcomes: the probability of having an unplanned readmission and the probability of entering institutional long-term care within the demonstration year. Our model specification for these outcomes was similar to that for continuous and count outcomes, but we used additional controls to account for differences between IAH and comparison group beneficiaries in the length of time they were exposed to the risk of the outcome (as a proportion of days eligible for the demonstration in a given demonstration year).

We measured outcomes for the period that beneficiaries were eligible during a given demonstration year, which started from the date of eligibility through the end of the demonstration year or date of death. Therefore, the eligibility period differed across beneficiaries, depending on their eligibility start dates and death dates. In particular, death occurred less frequently in each demonstration year for the IAH group than for the matched comparison group (for example, 19.4 percent of IAH beneficiaries died during Year 7 compared with 27.0 percent of matched comparison beneficiaries).[26] Such a difference implied the importance of controlling for observation length because, all else being equal, IAH beneficiaries spent more time during the demonstration year at risk for the outcome than did the matched comparison beneficiaries. Further, death might directly affect the probability of readmission (or entry into long-term care) if the probability changes as people approach death. Thus, not controlling for death could bias the estimated effect of the demonstration.

For continuous and count outcomes, we accounted for differential observation lengths by annualizing the outcome and using eligibility weights in regressions. However, because we could not annualize binary outcomes, we employed a modeling approach similar to the one used in Deb (2016). The basic idea behind Deb's model is to first estimate a survival model to derive the predicted probability of dying for each individual in each time period and then include the predicted

[26] Ideally, the matching process would result in a comparison group with the same expected survival (as of the eligibility date) as the IAH beneficiaries. But it is possible that factors not observable in claims data caused a differential expected survival between the two groups. To the extent that these differences are changing over time in a way that we cannot control for in claims data, the differential trend could lead to bias in our impact estimates.

probability of dying in the second stage to account for the differences in outcomes because of the differences in the death rate across individuals.

Following Deb's approach, we estimated a survival-adjusted difference-in-differences model, controlling for the predicted probability of dying within the demonstration year (h_{it}), the interaction between treatment status and the probability of dying, and the proportion of time during the demonstration year that the beneficiary was eligible and alive ($survdays_{it}$).

Equation (3) shows our model specification:

(3)

$$\begin{aligned} P(Y_{it}=1) = {} & \alpha + X_{it}\beta + \tau.treatment_{it} + \beta_h h_{it} + \beta_{Rh} treatment_i h_{it} + \beta_s survdays_{it} + \gamma_{-1}\, PD_1 \\ & + \gamma_1\, DY_1 + \gamma_2\, DY_2 + \gamma_3\, DY_3 + \gamma_4\, DY_4 + \gamma_5\, DY_5 + \gamma_6\, DY_6 + \gamma_7\, DY_7 + \theta_{-1} treatment_{it} \bullet PD_1 \\ & + \theta_1 treatment_{it} \bullet DY_1 + \theta_2 treatment_{it} \bullet DY_2 + \theta_3 treatment_{it} \bullet DY_3 + \theta_4 treatment_{it} \bullet DY_4 \\ & + \theta_5 treatment_{it} \bullet DY_5 + \theta_6 treatment_{it} \bullet DY_6 + \theta_7 treatment_{it} \bullet DY_7 + \omega_{it} \end{aligned}$$

Where Y_{it} is a binary variable for whether the beneficiary had an unplanned readmission (or long-term care entry); h_{it} is the predicted probability of dying in the demonstration year, derived from the estimated survival model in Equation (2); $survdays_{it}$ is the number of days from the beneficiary's eligibility date through the end of demonstration year or date of death, divided by 365 (or 366 for a leap year); and ω_{it} is a random error term. The remaining covariates are the same as those in Equation (1).

In Equation (3), the term h_{it} measures the predicted probability of beneficiaries dying in the year regardless of their actual survival or censoring status. Because h_{it} was derived from the difference-in-differences survival model, it accounted for any death rate difference between the IAH and comparison groups that was not captured in matching, as well as any death rate difference resulting from the demonstration. Coefficient β_h captured changes in the probability of readmission (long-term care entry) as the death rate increases, and coefficient β_{Rh} captured differential changes in this probability for those in the IAH group versus the comparison group. Last, coefficient β_s captured the effect of the length of time at risk of readmission (long-term care entry) conditional on predicted death.

We estimated Equation (3) using a logistic regression model. As with death, we adjusted standard errors for hybrid clustering and used matching weights to ensure

equal sizes of IAH and comparison groups. Because estimation of Equation (3) involves a generated regressor h_{it}, we bootstrapped our estimates and standard errors, employing a multiple-imputation approach (Deb 2016). After bootstrapping, we transformed $\theta_1 - \theta_7$ into difference-in-differences effect estimates, following steps similar to those we used in estimating impacts for other outcomes. For each outcome, we also estimated a separate difference-in-differences model that used a post-demonstration indicator and its interaction with the IAH status to obtain the five-year annual effect estimate.

6.3.4. Adjustment to standard errors for clustering

To obtain accurate estimates of standard errors for the impact estimates, it was important to account for possible clustering of observations within geographic areas. CMS selected certain practices to implement IAH, each of which serves beneficiaries in a specific area. We selected patients from the same geographic catchment area for the matched comparison group. The IAH group sample was clustered by practice in that geographic area—all beneficiaries who met the eligibility criteria and received home-based primary care from the same demonstration practice. We could not model practice-level clustering of the comparison group, however, because we selected those beneficiaries without knowledge of the practice from which they received their primary care. We accounted for this asymmetric clustering structure of the two groups in our regression to avoid overstating the precision of the estimates.

In addition to the practice-level clustering, we had multiple observations for some beneficiaries in the sample. Because the observations on a given beneficiary in one period clearly were not independent of the observations on the same beneficiary in other periods, our estimator of the variance had to account for this time dependence of repeated observations.

To account for asymmetric practice-level clustering and multiple observations for some beneficiaries, we used what we refer to as a hybrid clustering approach. This approach accounted for clustering at the practice level for the IAH group only and took into account the time dependence of repeated observations for IAH and comparison beneficiaries.[27] Implementing this approach meant that all IAH beneficiaries in a given site were from a single cluster. Because the entire practice was selected to provide IAH in the given area, we have to account for this clustering effect to avoid overstating the precision of the estimates (that is, to avoid standard errors that are too small, giving a false sense of confidence about the effect of the

[27] Accounting for clustering at the practice level for the treatment group captures the correlation among observations in each IAH practice, whether for the same person across time periods or different people in the same time period. We implemented the hybrid clustering approach in the statistical software used for the analysis (Stata) by defining a cluster variable that takes the value of the practice ID for the treatment group and the value of the beneficiary ID for the comparison group.

demonstration). To correctly identify the clustering effect in the IAH group, we did not include site fixed effects in the regression equation.[28]

Our approach to adjusting standard errors was consistent with the goal of evaluating only the practices that participated in the demonstration in this report. We could not generalize beyond the demonstration practices to home-based primary care provided across the nation as a whole because demonstration practices were not a random sample of all practices, and we did not know the extent to which IAH sites were similar to other practices and the types of patients they serve. Instead, we assumed that the IAH beneficiaries in a given practice were a random sample of all eligible beneficiaries of that practice. For this reason, our statistical tests accounted for the random variation among eligible beneficiaries who received care from the demonstration sites.

6.3.5. Weighting

Construction of beneficiary weights. For continuous and count outcomes, we estimated regressions with observations at the beneficiary level and weighted the observations to capture two factors: (1) the share of months a given beneficiary was eligible for the demonstration during each pre-demonstration or demonstration year and (2) the number of comparison beneficiaries matched to each treatment beneficiary. We referred to the former as the eligibility weight; it controlled for differences in the length of time that beneficiaries were observed during a given study year. We referred to the latter as the matching weight. Because we matched each treatment beneficiary to up to five comparison beneficiaries, applying matching weights ensured that the impact regression was not disproportionally weighted toward IAH beneficiaries who had more matched comparison beneficiaries (such as five versus two).

The construction of final beneficiary weights for continuous and count outcomes required three steps. First, we constructed the eligibility weight as the share of months eligible for the demonstration during each pre-demonstration or demonstration year. After we determined a beneficiary's eligibility for the demonstration in a given pre-demonstration or demonstration year, we included the beneficiary in the analysis sample beginning on the first day of the following month.

[28] Ideally, including site fixed effects would improve estimation by controlling for factors that varied across geographic areas and affected outcomes for IAH and comparison beneficiaries within a given area. But because all IAH beneficiaries in a given site (stratum) were from a single practice (cluster), controlling for stratification and clustering at the same level would lead to under-identification. That is, we could not identify the clustering effect with only one IAH group practice per site in a stratified design (Schochet 2008). Relative to the site fixed effects, clustering was by far the more important factor to account for when estimating the variance of the estimate. If we failed to account for clustering when estimating variance, the standard errors and statistical significance of the estimates would be misleading and could lead to incorrect conclusions about the impact of the demonstration. To avoid that problem, we could not take advantage of the gains we would have achieved by accounting for the stratified approach.

The beneficiary remained in our analysis sample for the entire year unless that beneficiary left Medicare FFS or died. For example, if a beneficiary entered the Year 7 sample on January 1, 2020, and died on June 20, 2020, that person was eligible for the demonstration for six months and thus had an eligibility weight of 0.5.

Second, we constructed matching weights to account for the size of the matched set. Each IAH beneficiary received a weight of 1, and each matched comparison beneficiary received a weight that was the inverse of the number of comparison beneficiaries within the matched set. For example, if an IAH beneficiary was matched to four comparison beneficiaries, each of the latter received a weight of 0.25. Comparison beneficiaries' matching weights ranged from 0.2 (if there were five matched comparisons for a particular IAH beneficiary) to 1 (one matched comparison). For all outcomes other than death, we obtained a composite weight by multiplying the eligibility weight by the matching weight.

In the third step, we created the final analytic weight for each beneficiary by rescaling the composite weight to ensure equality in the weighted number of IAH and comparison beneficiaries for each site and year.[29] As we described, we implemented hybrid clustering adjustments but could not use site fixed effects (an indicator for each site). Because beneficiaries had different eligibility weights, the number of weighted IAH beneficiaries in a given site and year might differ from the number of weighted comparison beneficiaries in the same site and year if we used the composite weight without rescaling it. For this reason, we rescaled the weights for comparison beneficiaries by site and year so that for each year, the weighted number of IAH beneficiaries equaled the weighted number of comparison group beneficiaries for each site. This approach ensured that the estimated treatment–comparison differences and the difference-in-differences estimates for each year accounted for any differential weighting of the IAH and comparison groups.

For binary outcomes, we used matching weights only. We did not include an eligibility weight in the death rate regression because the survival model we employed accounts for differential observation lengths for the outcome via a hazard function.[30] For other binary outcomes (probability of unplanned readmission and entering institutional long-term care), we used a survival-adjusted model for binary outcomes, which explicitly accounts for the effects of the death rate and time survived since eligibility. We describe the model specifications for these binary outcomes earlier in this section.

Practice-weight method. We refer to the above weighting scheme as beneficiary weighting. Under beneficiary weighting, large practices that served more

[29] For more information, refer to the evaluation report covering Years 1 to 4 of the IAH demonstration.

[30] The weights used for the survival regression did not have to be rescaled because, without any eligibility weights, the matching weights ensured that the weighted number of IAH and comparison beneficiaries for each site and year were equal to each other.

beneficiaries had more influence on the estimated effect and smaller practices had less influence. We also report estimation results based on an alternative weighting scheme that allows all practices to have equal influence on the estimated effect regardless of the size of their patient population. This method, which we refer to as the practice-weight method, yields an estimate that reflects the average effect of changes that practices made in response to the IAH payment incentive (and in Year 7, any changes in the relative effectiveness of care during the COVID-19 pandemic). Under this approach, each practice has equal influence on the results.

An example may help explain the difference between the two methods. Let us assume that the demonstration had 4 large practices of 1,000 patients each that did not change care delivery in response to the demonstration and 12 practices of 250 patients each that changed care delivery. The practice-weight method would give equal influence to all practices, and the estimate would reflect that most practices (12 out of 16) changed care delivery. The beneficiary-weight method, on the other hand, would give more influence to the 4 practices that served most of the beneficiaries (4,000 out of 7,000), and the estimate would reflect that most of the beneficiaries were treated by practices that did not change care delivery.

As the example illustrates, when the effect of the demonstration differs across IAH practices, the beneficiary-weight method would lead to an estimated effect that tends to resemble those of the largest practices, thus masking the effect on smaller practices. The beneficiary-weight estimate reported in Chapter 2 of the report can be informative for policymaking if the largest IAH practices are a representative sample of the largest home-based primary care providers in the nation. But although the IAH practices were selected to include diverse approaches to providing home-based primary care, they do not represent all practices in the nation that provide home-based primary care. Therefore, the practice-weight estimate provides an important understanding of the average effect of the demonstration across a variety of delivery models for home-based primary care. Considering beneficiary- and practice-weight methods together is important to our understanding of the effect of IAH on outcomes.

To implement the practice-weight method, we reweighted the final analytic weights for all IAH and matched comparison beneficiaries using a ratio that varied by site and year so that the summed weights among all beneficiaries in each practice were equal across all practices in each year (treating the three members of the Richmond-based consortium as separate sites). We then reestimated the effect on total Medicare expenditures as specified in Equation (1) using these new weights.

6.3.6. Control variables

Although our matching process ensured that the comparison groups were very similar to the IAH groups along many characteristics, there may still be important

differences in some of these characteristics that could affect the outcomes. Therefore, we included four types of control variables: (1) variables describing eligibility for the demonstration; (2) demographic characteristics; (3) ADL indicators; and (4) measures of health status, including HCC risk score, HCC indicators, and chronic condition indicators (Exhibit A.18). We included all HCC indicators and categories of HCCs used for matching (Exhibit A.12), but some of the other control variables were included in our regressions at a more detailed level than the variables we used in matching (to help ensure balance during matching). For example, we used three age categories in propensity score matching, whereas we used five age categories in the outcome regressions.

As we noted, we included a dummy variable for each year and an indicator of whether the beneficiary was in the IAH or comparison group. Because of the repeated cross-sections in our multiyear data set, we used contemporaneous control variables for all years of the demonstration; for example, in demonstration Year 7, we used the Year 7 values of all control variables whether or not a beneficiary appeared in the sample in an earlier demonstration year.

Exhibit A.18. Control variables used in regressions

Variable
Eligibility for the demonstration
Number of months since most recent hospital admission: 1, 2 or 3, 4 or more
Month of the demonstration year that beneficiary met eligibility criteria (1, 2 to 6, 7 to 12)[a]
Demographic characteristics
Age: younger than 65, 65 to 74, 75 to 79, 80 to 84, 85 or older
Gender
Race and ethnicity: White, Black, Hispanic, Asian, American Indian or Alaska Native, other, or unknown
Dually eligible for Medicare and Medicaid
Original reason for Medicare entitlement: old age, ESRD or ESRD and disability, disability only
ADLs
Number of ADLs for which beneficiary requires human assistance: 2, 3 or 4, 5 or 6
Whether information about the feeding ADL was missing[b]
Health status[c]
HCC risk score
Specific HCCs
HCC 8, metastatic cancer and acute leukemia
HCC 9–10, lung and other severe cancers; lymphoma and other cancers
HCC 11–12, colorectal, bladder, and other cancers; breast, prostate, and other cancers and tumors
HCC 18, diabetes with chronic complications
HCC 21, protein-calorie malnutrition
HCC 27, end-stage liver disease
HCC 28–29, cirrhosis of liver; chronic hepatitis
HCC 46, severe hematological disorders
HCC 48, coagulation defects and other specified hematological disorders
HCC 51, dementia with complications
HCC 52, dementia without complications
HCC 54–55, drug/alcohol psychosis; drug/alcohol dependence
HCC 57–58, schizophrenia; major depressive, bipolar, and paranoid disorders
HCC 70–71, quadriplegia; paraplegia
HCC 72, spinal cord disorders/injuries
HCC 85, congestive heart failure
HCC 96, specified heart arrhythmias
HCC 103–104, hemiplegia/hemiparesis; monoplegia, other paralytic syndromes
HCC 106, atherosclerosis of the extremities with ulceration or gangrene
HCC 107–108, vascular disease with complications; vascular disease
HCC 111, chronic obstructive pulmonary disease

Exhibit A.18 (*continued*)

Variable
HCC 134, dialysis status
HCC 136–138, chronic kidney disease, stages 3–5
HCC 139–140, chronic kidney disease, stages 1–2 or unspecified; unspecified renal failure
HCC 157–159, pressure ulcer of skin with necrosis through to muscle, tendon, or bone; or with full or partial thickness skin loss
Chronic conditions measured by Chronic Conditions Data Warehouse
Alzheimer's disease or related disorders
Acute myocardial infarction or ischemic heart disease
Asthma
Hip or pelvic fracture
Stroke or transient ischemic attack
Number of chronic conditions and the square of the number of conditions
Other measures of health status
Anemia[d]
Fluid and electrolyte disorders[d]
Chronically critically ill or medically complex diagnosis

[a] For all sites in Year 6, month 1 was January. For all sites in Years 4 and 5, month 1 was October. In Years 1 to 3, sites began the demonstration in June or September each year. For sites that began in June, month 1 is June. For sites that began in September, month 1 was September.

[b] Feeding assessments were not available with home health assessment data at the time of recertification. If the beneficiary had a previous assessment during the study year that was recorded at the time of discharge from home health care, we used the feeding values from that assessment. Sometimes, however, there was no previous discharge assessment.

[c] For binary outcomes (death, entry into institutional long-term care, and readmission), we estimated a survival-adjusted difference-in-differences model which controlled for additional variables: the predicted probability of dying within the demonstration year, the interaction between treatment status and the probability of dying, and the proportion of time during the demonstration year that the beneficiary was eligible and alive.

[d] Measured using claims from the most recent inpatient stay and observation stay in the year before the demonstration eligibility date. We drew diagnosis codes for these conditions from Gagne et al. (2011).

ADL = activity of daily living; ESRD = end-stage renal disease; HCC = Hierarchical Condition Category.

6.3.7. Testing the parallel trends assumption

The validity of the difference-in-differences estimates as impact estimates of the demonstration rely on the classic difference-in-differences parallel trends assumption that there was no significant differential trend between the IAH and matched comparison groups during the pre-demonstration period. That is, outcomes should change at the same rate for both groups in the two-year pre-demonstration period, so any difference in outcomes between the two groups would remain the same during that period. Therefore, the difference-in-differences estimate for two years before the demonstration, θ_{-1}, served two purposes: (1) it ruled out or identified

treatment-comparison differences in trends during the pre-demonstration period and (2) in so doing, it helped inform the more important difference-in-differences analysis for the demonstration period. Specifically, a statistically significant θ_{-1} would indicate that the difference in a given outcome between the IAH and comparison groups changed significantly from two years before the demonstration to the year before the demonstration. In this case, the IAH and comparison groups could have been on nonparallel outcome trajectories during the pre-demonstration period. We referred to nonparallel outcome trajectories during the pre-demonstration period as a pre-existing difference in trend.

The possible presence of nonparallel pre-demonstration trends would have limited our confidence in the demonstration impact estimates for a given outcome. This was because the difference-in-differences estimates for the demonstration years could have reflected the continuation of a pattern (for example, narrowing or widening differences between the two groups) that began during the pre-demonstration period rather than reflecting an impact of the demonstration.

We examined the difference-in-differences estimate for two years before the demonstration for all outcomes reported. The estimate was not statistically significant and was small as a percentage for most outcomes, including total expenditures and hospital care use, suggesting that the parallel-trend assumption held for those outcomes between those two years. In the Year 7 sample, we found statistically significant difference-in-differences estimates for two years before the demonstration for hospice spending, death, and the probability of entering institutional long-term care, which violated the parallel-trend assumption. In both cases, the estimated trend in the pre-demonstration period was relatively large compared with the pre-demonstration IAH group mean outcome. Because it was impossible to rule out the possibility of truly nonparallel pre-existing trends for these outcomes in which the difference-in-differences estimate for two years before the demonstration was different from zero, we were cautious in interpreting the impact estimates for these outcomes.[31]

6.3.8. Assessing the relative influence of individual practices

As we noted, under beneficiary weighting, practices have different amounts of influence on the estimated effect depending on their sizes. To understand which practices drove the estimated effects of the demonstration, we reestimated the

[31] It is possible to control for pre-existing trends by including linear time trends in the regression, but this approach would impose an overly restrictive assumption on our model: that the one-year pre-demonstration trends would continue throughout the demonstration.

beneficiary-weight regression, leaving out one practice at a time.[32] Specifically, we estimated 12 regressions (treating each member of the Richmond-based consortium separately), with each regression excluding the IAH beneficiaries from one practice and their matched comparisons in all years. If all 12 regressions showed similar estimates for Year 7 as the main regression, we would conclude that all practices equally influenced the full sample estimate. On the other hand, if excluding a given practice substantially changed the estimated effect, we would conclude that the site strongly influenced the full sample estimate. In Appendix B, we report the Year 7 estimated effects from each of the 12 regressions for total expenditures.

6.3.9. Average annual effect of IAH on Medicare spending through the first seven years

In addition to estimating the yearly effect of IAH on total Medicare spending, we estimated the average annual effect through the first seven years of the demonstration. Our estimates of the yearly effects of the demonstration use separate regression models (with different samples) to estimate effects in Years 1 to 5, Year 6, and Year 7. Combining these estimated effects from individual years to create an estimated average effect across the first seven years has conceptual and technical limitations.

Conceptually, combining effects from different years with different samples can lead to inferences based on an average that does not represent the actual experience of the IAH demonstration in any year and has no generalizability to other years and other home-based primary care practices. This is particularly true in this case for two reasons. First, in Year 7, the COVID-19 pandemic makes it difficult to estimate the extent to which the IAH demonstration contributed to the estimated effects (see additional explanation in Chapter 4 of the report). If effects were different in Year 7 because of the pandemic, combining Year 7 with pre-pandemic years would create an average estimate that is difficult to interpret. Second, our estimates of the effect of IAH on outcomes are based on three samples of IAH practices from separate regressions, as described in Section 6.1 of this Appendix. As we described in the annual report covering IAH Years 1 to 5, one practice which left the demonstration after Year 5 and no longer provides home-based primary care greatly affected estimated effects while it was in the demonstration. Therefore, combining estimates from Years 1 to 5 of the demonstration with estimates from Years 6 and 7 mixes samples in a way that affects interpretation.

Technically, combining the yearly effect estimates using an average weighted by beneficiary-month observations is straightforward. However, combining the

[32] We could have estimated regressions separately for each site to obtain site-specific estimated effects, but the statistical power for some of these regressions was too low because of the small sample sizes at the site level. Estimating the regressions and excluding one site at a time enabled us to assess the influence of each site by comparing those estimates with the estimate from the full sample.

estimates of precision (that is, standard errors and confidence intervals) is more complex. Since the estimates come from different regression models, we must account for repeat observations across models at the site and the beneficiary levels to generate a confidence interval around the estimated average annual effect. Failing to do so can lead to confidence intervals that are too narrow.

To deal with this technical limitation, we used seemingly unrelated regression, a post-estimation regression technique that combines parameters and variance-covariance matrices from different models into a single, simultaneous, and sandwich robust covariance matrix. From that combined matrix, we can test hypotheses, such as whether the linear combination of all the parameters estimating yearly effects of IAH, averaged, is different from zero. Since this method does not explicitly model the correlation between models arising from overlapping sites and beneficiaries, it may have understated the true correlation between estimates, leading to confidence intervals that were too narrow. In other words, this approach could have led to estimated effects that were more precise than they should have been. Still, it is a straightforward way to combine estimates across years that makes some correction for repeated observations.

To estimate average annual effects, we took the following steps:

1. We re-estimated models estimating the yearly effects of IAH on total Medicare spending, saving their coefficients and variance-covariance matrices for each of three models with different samples: (1) Years 1 to 5 with all IAH sites; (2) Year 6 using only sites that participated in Years 1 to 6; and (3) Year 7 using only sites that participated in Years 1 to 7. Those models produce estimates that correspond to the reported effects of IAH in each year (as shown in Exhibit 3.1 in Chapter 3 of the report).
2. We used the Stata command for seemingly unrelated estimation to combine the variance-covariance matrix of the models using cluster robust standard errors.
3. We used linear combination commands and a Wald test to test whether the average of the yearly effect estimates weighted by beneficiary person months from Years 1 to 5, Year 6, and Year 7 was different from zero.

Because we were concerned conceptually about combining estimates from different IAH samples, we also estimated an average annual effect dropping one site at a time in all years.

6.3.10. Sensitivity analyses controlling for COVID-19 diagnosis and hospitalization

Given the importance of the COVID-19 pandemic and its effect on health care, we conducted sensitivity analyses in Year 7 (2020, the first year of the pandemic) that controlled for COVID-19 diagnosis and COVID-19 hospitalization in our regressions. Large differences in models that do and do not control for COVID-19 diagnosis or

hospitalization would suggest a strong confounding effect of COVID-19 on the relationship between IAH participation and outcomes that might cause us to interpret our main results differently. To test the influence of COVID-19 diagnosis or hospitalization, we created two binary indicators of COVID-19 exposure: (1) diagnosis with COVID-19 at any point in calendar year 2020 (identified using diagnosis codes in the clinician/supplier, outpatient, inpatient, and SNF claims) and (2) any hospital admission or observation stay with a diagnosis code for COVID-19 between the date the beneficiary became eligible for IAH and entered our sample and the end of the demonstration year. We then included, in separate models, the binary indicators for COVID-19 diagnosis and hospitalization in regressions for total Medicare spending.

6.3.11. Subgroup analysis

To better understand the large, estimated effects in Year 7, we estimated effects for subgroups of beneficiaries who might have been at higher risk of poor outcomes and high spending as a result of social and health care disruptions during the pandemic than other beneficiaries in our sample. We defined the following subgroups for analyses: (1) beneficiaries needing assistance from another person with more ADLs (5 or all 6) versus beneficiaries needing assistance with fewer ADLs (2 to 4); (2) beneficiaries ages 85 and older versus beneficiaries ages 84 and younger; (3) beneficiaries with 10 or more chronic conditions versus beneficiaries with less than 10 chronic conditions; (4) beneficiaries with dual Medicare and Medicaid coverage versus beneficiaries with Medicare alone; (5) beneficiaries whose original reason for entitlement was based on disability, end-stage renal disease, or both versus beneficiaries whose original reason for entitlement was based on age; and (6) non-Hispanic White beneficiaries versus either non-Hispanic Black beneficiaries or other beneficiaries.

To estimate effects for these subgroups, we ran regressions for total Medicare spending and inpatient spending that interacted binary subgroup indicators based on the definitions above with all covariates (including IAH participation and year). Estimating models this way (instead of an equivalent stratification approach) allowed us to use a Wald test to determine whether the estimated effects of IAH in Year 7 for one group (for example, beneficiaries aged 85 and older) were different from the other group (for example beneficiaries aged 84 and younger). We found statistically significant differences in the estimated effects of IAH by group only in the 5 to 6 ADL versus 2 to 4 ADL subgroup (Exhibit 2.2 and Appendix B).

6.4. Bayesian difference-in-differences model

6.4.1. Overview

In addition to the frequentist (traditional) analyses we described earlier in this chapter, we conducted a set of analyses using the Bayesian statistical paradigm. Assessing the effects of IAH probabilistically, as Bayesian techniques permit, maintains a rigorous statistical standard and provides a more flexible interpretation of the program's effects. The frequentist approach classifies the demonstration's impact as statistically significant or not statistically significant; in contrast, a Bayesian analysis allows probabilistic estimates about whether the demonstration achieved a certain outcome. For example, one could conclude that "there was a 98 percent chance that the IAH demonstration incentive reduced expenditures by at least $100 PBPM in demonstration Year 7." Such conclusions offer the opportunity to tailor inferences to substantive questions of interest and apply subject matter expertise in deeming meaningful effects.

Overall, the Bayesian and frequentist analyses were similar, but they had some differences. As with the frequentist approach, the Bayesian analysis used a comparison group difference-in-differences design to identify effects attributable to the IAH demonstration. The outcome of interest was total Medicare expenditures PBPM. We used the same data sets for the frequentist and Bayesian analyses. Moreover, we used the same eligibility and matching weights and the same control variables. The Bayesian analysis diverged from the frequentist analysis, however, in three ways. In this chapter, we describe the three factors that differentiated the Bayesian analyses from their frequentist counterparts: the prior distributions, the method used to account for clustering, and the computational approach used to fit the models.

Prior distribution. Assigning a prior distribution to each model parameter translated the model into the Bayesian framework and allowed for probabilistic inference. We placed a standard normal prior distribution—denoted $N(0,1)$ — on the overall impact of IAH. By doing so, we incorporated a prior expectation that very large positive or negative impacts of IAH on expenditures were substantially less likely than small and moderate impacts. We based our prior expectation on the general result that other evaluations of the impact of home-based primary care and other interventions for chronically ill, frail beneficiaries very rarely show effect sizes larger than two standard deviations. We centered the normal distribution at a mean of zero to remain agnostic about whether the IAH demonstration would be successful.

Method used to account for clustering. The full Bayesian model accounted for clustering by using random effects, and the frequentist analysis used cluster-robust standard errors (as we described earlier in this appendix). Specifically, the two-stage

full Bayesian model accounted for clustering using beneficiary- and site-specific random effects for the IAH and comparison groups, in which each site included IAH beneficiaries from a demonstration practice and their matched comparison beneficiaries. In contrast, the frequentist analysis estimated cluster-robust standard errors, which assumed that IAH beneficiaries were clustered by practices and comparison beneficiaries were clustered by individual beneficiaries rather than practices (a hybrid clustering approach). The Bayesian model could not apply the same approach because it accounted for clustering using random effects instead of cluster-robust standard errors.[33] This methodological difference in accounting for clustering could lead to differences in point estimate and standard error of the estimate.

Two-stage model. We further modified the frequentist model to make Bayesian computationally feasible. We adopted these modifications purely as a computational convenience; they are not inherently Bayesian, and a traditional impact estimation framework could also adopt this approach. Ideally, we would have liked to fit a single, unified model at the beneficiary level, as in the frequentist analysis, but such a model would have taken a prohibitively long time to converge on our analysis platform. Because of time constraints, we used a two-stage approximation of this ideal beneficiary-level model. In the first stage, we aggregated the beneficiary-level data set to the site level. Using output from Stage 1, we estimated the impact of the IAH demonstration using a Bayesian difference-in-differences framework in Stage 2.

6.4.2. Full Bayesian model, pooled

To understand the full Bayesian model, we begin by presenting a single unified model at the beneficiary level. As we show in Equation (4), this procedure accomplishes impact estimation and risk adjustment simultaneously through a model of the following form:

(4) $$Y_{ijt} = \alpha + X_{it}\beta + \tau z_{it} + \gamma_t + \theta_t z_{it} + a_i + b_j + c_j z_{it} + d_{jzt} + \varepsilon_{it}$$

This model uses a slightly different notation than its frequentist counterpart, Equation (1), for clarity of presentation of the random effects.

- We use i to index beneficiaries; $j = 1,\ldots,12$ to index geographic areas (or, loosely speaking, sites at which both IAH and comparison beneficiaries resided in Year 7); and $t = -1,\ldots,7$ to index years.

[33] A Bayesian model requires a fully model-based approach to account for clustering, and cluster-robust standard errors are an adjustment performed after the modeling process.

- Y_{ijt} is total Medicare expenditures PBPM measured for beneficiary *i* from site *j* in year *t*; X_{it} is a set of beneficiary characteristics measured in year *t*; z_{it} is the treatment status of beneficiary *i* in year *t*.
- Greek letters denote parameters to be estimated: α is a constant term; β contains the effects of the beneficiary characteristics; τ captures any differences between IAH and comparison beneficiaries in the year before the demonstration that persist despite matching; γ describes the secular time trend that applies to both IAH and comparison beneficiaries; and the θs are the difference-in-differences impacts of interest. As with the frequentist model, we estimated γ_{-1} and θ_{-1} for two years before the demonstration, and $\gamma_1 - \gamma_7$ and $\theta_1 - \theta_7$ for each of the six demonstration years. Note that $t = 0$ corresponds to the baseline year (the year before the demonstration), so we omitted γ_0 and θ_0 from the model.
- Random effects are denoted by Roman letters: the as and bs are beneficiary- and site-level random intercepts, respectively, which account for the correlation across repeated observations on a given beneficiary or site; the cs are site-specific baseline IAH/comparison differences; and the ds are site-treat-year random intercepts. We assume that the as and ds each follow a univariate normal distribution, and the bs and cs jointly follow a bivariate normal distribution. The latter assumption allowed for correlation between a site's intercept and the IAH/comparison difference in that site.

Last, we weighted the regression using the same weighting scheme (beneficiary weighting) that we used in the frequentist analysis, as we discussed earlier.

We calculated the adjusted total Medicare expenditures for the IAH and matched comparison groups in each year, the difference-in-differences estimates (θ_{-1}, $\theta_1 - \theta_7$), and percentage impact relative to unadjusted IAH group mean expenditures in the year before the demonstration. In addition, we estimated the probability of reducing expenditures by at least $100 PBPM. In all calculations, we adjusted the yearly average outcomes for both groups to reflect the covariate distribution of the IAH group in the latest (seventh) demonstration year—the same approach we used in the frequentist analysis.

Because of the number of observations in the data set, fitting Equation (4) as a single, unified model at the beneficiary level was computationally prohibitive. For this reason, we fitted the full Bayesian model using a two-stage approximation to decrease computational run times. The first-stage model was a beneficiary-year-level risk adjustment fit using hierarchical linear regression. The goals of the first-stage analysis were to aggregate beneficiaries to the site level and risk-adjust outcomes to

enable comparisons across sites and years whose case mix differed (Equation [5]). In the first-stage model, we adjusted for the same beneficiary-level covariates as the frequentist model (see Exhibit A.18). We used the risk-adjusted site-year-level output from Stage 1 as data in Stage 2, which estimated the impact of IAH demonstration in a Bayesian difference-in-differences framework (Equation [6]).

(5) Stage 1: $Y_{ijt} = A_{jz_{it}t} + X_{it}\beta + a_i + \varepsilon_{it}$

As we described, the site-treatment-year effect A_{jzt} represents the estimated fixed effect for site *j* and treatment group *z* in year *t*. There were 216 such fixed effects from two groups (IAH and comparison) from each of the 12 sites in total (counting each member of the consortium as a separate site). The parameters β describe the effects of beneficiary-level control variables X_{it}, whereas beneficiary-level random effects a_i account for correlations across repeated observations on beneficiary *i*. We assumed that the beneficiary-level random effects a_i and the overall error term ε_{it} came from a normal distribution with mean zero and its own variance. Similar to the frequentist model, we used the rescaled composite weights for the Stage 1 model. Then, we used the aggregated site-treatment-year estimates ($\hat{A}_{jzt}$) and associated standard errors (s_{zjt}) from the Stage 1 model when we estimated the Stage 2 full Bayesian difference-in-differences regression (Equation [6]).

(6) Stage 2: $\hat{A}_{jzt} = \alpha + \tau z + \gamma_t + \theta_t z + b_j + c_j z + d_{jzt} + \varepsilon_{jzt}$

In the Stage 2 model, we included an overall intercept α and controls for the secular time trend γ_t and treatment τ. We accounted for clustering through random effects b_j, c_j, and d_{jzt}, as we described earlier. The parameters of interest, θ_t, represent the overall difference-in-differences terms.

We assigned a standard normal distribution— $Normal(0,1)$ —as the prior for each model parameter: $\alpha \sim N(0,1)$, $\tau \sim N(0,1)$, $\gamma \sim N(0,1)$, $\theta \sim N(0,1)$, $(b_j, c_j) \sim MVN(0,\Sigma)$, $d \sim N(0,\sigma^2)$, where σ^2 is the overall noise variance. The prior for Σ included two parts: one part to address correlations between b_j and c_j, and one to address the standard deviation of b_j and c_j. The former part took on an LKJ correlation prior (Lewandowski et al. 2009); the latter took on a standard normal distribution. The multiplication of these two parts constituted the prior on $\Sigma : \Sigma = \begin{pmatrix} \sigma_b & 0 \\ 0 & \sigma_c \end{pmatrix} \Omega \begin{pmatrix} \sigma_b & 0 \\ 0 & \sigma_c \end{pmatrix}$ where $\sigma_b, \sigma_c \sim N(0,1)$ and $\Omega \sim \mathrm{LKJ}(2)$.[34] Last, our prior on the error term is given by $\varepsilon_{jzt} \sim Normal(0, s^2_{jzt})$. Therefore, both σ^2 and s^2_{jzt} act as weights in Stage 2. We used the "lme4" package in R to fit the Stage 1 model. For Stage 2, we used a novel probabilistic programming language called Stan, which provides fast, full Bayesian inference, even for complex models.

[34] LKJ is a distribution on correlation matrices (usually called Ω). The distribution has one parameter, ν, so $\Omega \sim$ LKJ(ν). When $\nu = 1$, the distribution is uniform over all possible correlation matrices. As ν increases, the distribution is more concentrated on the identity matrix, which corresponds to zero correlations. Thus, for $\nu = 2$, the distribution slightly favors less correlation, shrinking the correlations somewhat toward zero. This is a weakly informative prior that helps stabilize the estimation.

7. Accounting for other programs and payments

Other programs administrated by CMS that take place concurrently with the IAH demonstration could influence our estimates of the effect of IAH on total Medicare expenditures. We addressed the potential influence of two CMS programs that providers for IAH or comparison beneficiaries may have participated in during 2019: the Quality Payment Program and ACOs.

7.1. Excluding Quality Payment Program payment adjustments in estimating the impact of the IAH

7.1.1. Background

As required by law, in 2017, CMS implemented the Quality Payment Program, which aims to incentivize clinicians to provide high-value care. Clinicians can participate in two tracks in the model: the Merit-based Incentive Payment System (MIPS) and alternative payment models (APMs). CMS pays clinicians for performance, participation, or both in these programs. Though performance measurement began in 2017, the first year it is relevant to estimating impacts for IAH is 2019 because that is the first year CMS made payments under the Quality Payment Program. This coincided with Year 6 for IAH, which began January 1, 2019, and continued into Year 7 (2020). In the first year of issuing payments, payments were small, with a maximum adjustment of just 1.9 percent (Navathe et al. 2019). Still, these adjustments could affect the estimate of IAH on total Medicare expenditures if IAH practices received larger (or smaller) MIPS adjustments than did comparison practices.

The MIPS program evaluates qualifying clinicians' performance across four domains and, based on performance, adjusts payments for professional services covered under the Medicare Physician Fee Schedule. The program is not voluntary; all MIPS-eligible clinician types who meet a threshold for volume of services are subject to a MIPS payment adjustment. CMS applies MIPS adjustments to payments for clinicians' professional services claims, and claims-based measures of expenditures include these adjustments by default.[35]

The APM track of the Quality Payment Program allows clinicians to participate in customized payment models that often seek to engage specific populations or care delivery approaches. Examples include Comprehensive Primary Care Plus, IAH, and Bundled Payments for Care Improvement models. Participation in an APM can offer additional bonus payments and changes the MIPS reporting requirements. CMS distinguishes between two types of APMs: Advanced APMs and MIPS APMs.

[35] Though adjustments in 2020 were made based on performance in 2018, we account for Medicare expenditures in the year they were made, so we began to count MIPS adjustments in 2020.

- Most Advanced APMs require participants to bear significant financial risk that is determined on an individual basis for each Advanced APM (for example, by setting an expenditure target above which the Advanced APM is responsible for costs). Clinicians who achieve threshold levels of patients or payments through the Advanced APM are eligible to receive a 5 percent APM incentive payment on their professional services claims and are excluded from MIPS and its reporting requirements. Unlike MIPS adjustments, the 5 percent incentive payments paid to Advanced APM participants are *not* included as part of claims and are paid separately to qualifying clinicians.
- MIPS APMs refer to APMs not designated as Advanced and include clinicians eligible for MIPS. Unlike Advanced APMs, clinicians participating in a MIPS APM are not eligible for the 5 percent incentive payment on professional services claims and are required to participate in MIPS and its reporting requirements. Clinicians participating in a MIPS APM are still subject to MIPS adjustments (like their peers nationwide not participating in any APM) but with modified performance category weighting and reporting requirements. For example, in the 2020 performance year, the cost performance domain received zero weight for MIPS APM participants (compared with 15 percent for non-MIPS APM participants). IAH is a MIPS APM.

7.1.2. Payments included when estimating impacts of the IAH demonstration

We included MIPS adjustments on payments made to MIPS-eligible clinicians in 2020 in our primary analysis of total Medicare expenditures. By default, the professional services claims include MIPS adjustments. Including these payments in our measure of total Medicare expenditures more accurately reflected the total amount CMS paid for services in 2020 compared with not including them. We also conducted an additional analysis excluding these payments to see whether they changed the Medicare expenditure outcomes in our sample in a way that influenced our estimate of the effect of IAH.

We did not include Advanced APM payments made to clinicians in 2020 in any analysis of total Medicare expenditures. Because these payments are calculated at the clinician level and are not reported in Medicare claims data, we would have to assign a portion of a clinician's payments to beneficiaries based on assumptions and data analysis. We do not believe this approach would be appropriate for our sample. First, because we do not attribute comparison beneficiaries to specific clinicians (or practices), we do not know which clinicians that received an Advanced APM payment provided care for a given comparison beneficiary. Second, because a clinician who had at least one patient in the comparison group likely had many patients who were *not* included in the comparison group, we do not know the share of that clinician's payment that we ought to assign to a single comparison beneficiary per month.

If Advanced APM payments were relatively large or if the share of IAH beneficiaries seen by clinicians in an Advanced APM was substantially different from the share of comparison beneficiaries seen by Advanced APM clinicians, then we may have misrepresented total Medicare expenditures and therefore the impact of IAH. These payments, however, probably would not have counted for a substantial component of total Medicare expenditures for beneficiaries in our sample. Advanced APM payments are most likely to make a difference to a measure of total expenditures (1) if the sample includes all FFS beneficiaries treated by a particular practice or clinician and (2) if a fairly large share of total expenditures for those beneficiaries were for services under the Physician Fee Schedule. Neither of these is the case for the IAH evaluation:

- Participation in Advanced APMs is substantially lower than MIPS. Fewer than 100,000 clinicians (of the approximately 1 million nationwide) received an Advanced APM payment in 2019.[36] Therefore, any effects of additional payments will likely (although we cannot be certain) be limited to a small share of comparison or IAH beneficiaries, who themselves represent only a small share of an individual practice's total FFS patient population.
- Total expenditures for IAH-eligible beneficiaries—which include the IAH and comparison groups—are less likely to be substantially affected by Advanced APM bonus payments made to clinicians compared with the general FFS Medicare population. Only relatively high-cost beneficiaries meet IAH eligibility requirements, and among this group, only a small share of total expenditures is for services that appear in clinician/supplier claims. Specifically, in Year 7, clinician/supplier claims accounted for only 16 percent of total expenditures for IAH beneficiaries and 15 percent for comparison beneficiaries, and only a minority of these would have been claims for services provided by clinicians who received an Advanced APM payment. These services represent a larger share of total expenditures for the general FFS Medicare population because the general population tends to have lower expenditures for inpatient, skilled nursing, and other services than the IAH-eligible population.

[36] CMS has not published information on participating Advanced APM clinicians in 2020. However, the list of qualifying initiatives is the same as 2019, and as such, we anticipate participation to be largely similar.

7.1.3. Excluding MIPS payment adjustments

To assess whether MIPS payments affected Medicare expenditures for beneficiaries in our sample in a way that affected the estimated impact of IAH, we repeated our total expenditure regression models after removing MIPS payment adjustments from the expenditures measure. MIPS adjustment amounts appear directly in the clinician/supplier and outpatient claim files. We removed the sequestration amount before subtracting the MIPS adjustment and then reapplied sequestration.[37]

7.2. Controlling for ACO participation in estimating the impact of IAH

7.2.1. Background

During the IAH demonstration, participation in ACOs increased. ACOs are groups of clinicians, hospitals, and other health care providers held accountable for the quality and cost of their patients' care. If an ACO achieves savings for CMS relative to a target spending level, then the ACO can share in those savings if it meets certain quality measures. Created as part of the Affordable Care Act, ACOs have become more common over time. By 2019, there were 487 ACOs in the Medicare program with 10.4 million assigned Medicare beneficiaries, up from 220 organizations and 3.2 million beneficiaries at the start of 2013 (CMS 2021a). Most of the IAH practices joined an ACO during the demonstration, and by Year 7, over four-fifths of IAH practices (representing 82 percent of IAH beneficiaries) and two-fifths of comparison beneficiaries participated in an ACO.

Simultaneous participation in IAH and an ACO could affect estimates of the impact of IAH. If IAH practices that were in ACOs changed their care delivery because of their ACO involvement, those changes could have affected their patients' Medicare expenditures (perhaps by improvements in care coordination that reduce total Medicare expenditures). ACO participation may also affect other outcomes of interest such as use and quality of care. For example, three IAH practices reported to us in Year 6 that they had worked with their ACO to implement a uniform process aimed at reducing hospital use among patients with a urinary tract infection or chronic obstructive pulmonary disease. If there were a significantly different effect of ACO participation on IAH practices relative to the comparison group, the estimated impact of the demonstration would reflect a combined impact of IAH and participation in the ACO.

To assess the influence of ACO participation on the estimated impact of IAH, we categorized beneficiaries as participating or not participating (or more specifically, being attributed to a provider which is or is not participating) in an ACO in each pre-

[37] Because CMS applies sequestration to payments as the last step in generating the total paid amount, to remove MIPS payments, we must first undo sequestration by multiplying claims by 1.02, then remove the MIPS adjustment, and then reapply sequestration to the final amount.

demonstration and demonstration year and controlled for this participation in regression analyses. We measured participation in the ACO using attribution for payment calculations by CMS (which we refer to simply as attribution in this chapter) or additional information gathered directly from the IAH practices, as we describe next.[38]

7.2.2. Assigning ACO participation to IAH and comparison beneficiaries

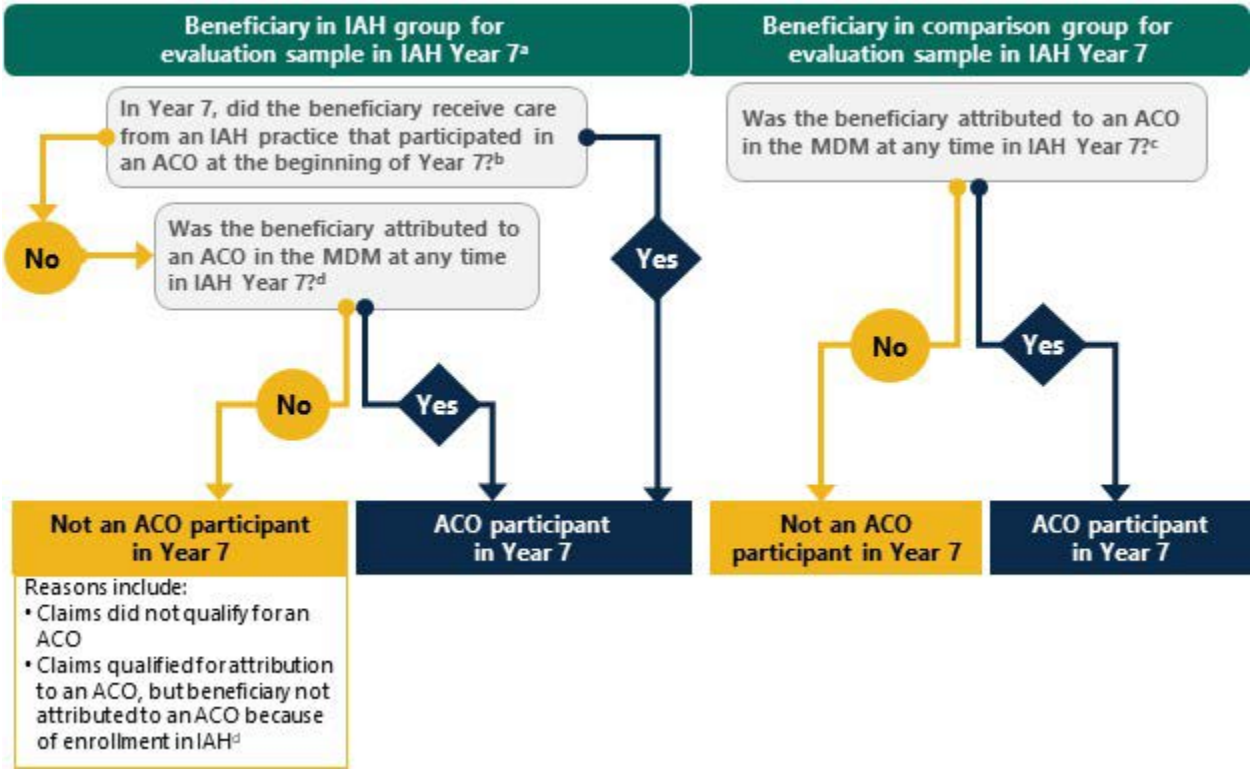
A key challenge to assessing the influence of ACO participation on the estimated impact of IAH is to identify beneficiaries in our sample reached by ACOs. Attribution to an ACO initiative depends on having qualified claims from ACO providers as measured by Medicare claims that have the National Provider Identifier and Tax Identifier Number (TIN) of a clinician at an organization that participates in the ACO initiative. CMS stores beneficiary-level data on attribution to an ACO initiative as well as participation in other initiatives such as IAH in the CMS Master Data Management (MDM) system, which is a longitudinal database system.

By design, beneficiaries are generally attributed to a single initiative in the MDM at a time (to avoid CMS making multiple payments for the same beneficiary at the same time). Beneficiaries enrolled in IAH are identified as such in the MDM during their period of enrollment. Enrollment in IAH in the MDM takes precedence over attribution in the MDM to any other initiative, including ACO initiatives.[39] Beneficiaries in the IAH evaluation sample but not enrolled in the IAH demonstration (discussed in Chapter 3 of this appendix) are not identified as participating in IAH in the MDM; therefore, these beneficiaries are eligible for attribution to an ACO in the MDM. Because enrollment in IAH in the MDM takes precedence, we supplemented MDM attribution data with qualitative information gathered from the IAH sites on ACO participation. After identifying which IAH practices were part of an ACO, we assigned IAH beneficiaries to an ACO based on the ACO participation of the practice from where they received care in a given demonstration year (Exhibit A.19).

[38] Throughout this analysis, we specify two distinct terms related to beneficiary involvement with ACOs. Participation denotes a beneficiary's relationship with an ACO for the purpose of our analysis. Designation of participation is based on information from CMS as well as qualitative information collected from sites. Attribution refers to CMS's formal designation related to identifying the population of beneficiaries relevant when calculating payments for ACOs. Beneficiaries are considered attributed to an ACO based entirely on their status in the CMS Master Data Management (MDM) system.

[39] It is possible for a beneficiary to be simultaneously enrolled in IAH and attributed to an ACO in the MDM. Our analysis accounts for this possibility and identifies those beneficiaries as participating in an ACO. This situation is, however, rare; the vast majority of beneficiaries enrolled in IAH according to the MDM would not be simultaneously attributed to an ACO because of the rule that IAH takes precedence. Although IAH practices could be part of ACOs and IAH beneficiaries could be treated by providers in these or other ACOs, CMS did not allow IAH beneficiaries to be attributed to an ACO for purposes of calculating ACO savings and payments.

Exhibit A.19. ACO assignment for IAH and comparison group beneficiaries



Note: IAH Year 7 was January to December 2020.

[a] We identified the IAH evaluation group using Medicare claims and other administrative data, and it included beneficiaries who were enrolled in the demonstration and some who were not enrolled. See Chapter 3 of this appendix for details on the differences between the IAH enrolled and evaluation samples.

[b] We based measurement of ACO participation for IAH practices largely on qualitative information provided by the IAH practices to Mathematica.

[c] CMS stores beneficiary-level data on IAH enrollment, attribution to an ACO initiative, and participation in other initiatives in the MDM system.

[d] Beneficiaries enrolled in IAH are identified as such in the MDM during their period of enrollment. Enrollment in IAH in the MDM takes precedence over attribution in the MDM to any other initiative, including ACO initiatives. Beneficiaries who are in the IAH evaluation sample but not enrolled in the IAH demonstration are not identified as participating in IAH in the MDM; therefore, these beneficiaries are eligible for attribution to an ACO in the MDM.

ACO = accountable care organization; MDM = Master Data Management system.

We identified IAH beneficiaries who participated in an ACO in each study year (demonstration Years 1 to 7 and two pre-demonstration years) based on the following logic:[40]

- We assigned IAH beneficiaries to an ACO if their IAH practice TIN was participating in an ACO at the beginning of the study year.[41]

[40] IAH enrollment in the MDM does not align perfectly with inclusion in the IAH group we identified for the evaluation. See Chapter 3 of this appendix for a detailed explanation of the difference between the IAH group used for the evaluation and the list of IAH enrollees identified by the implementation contractor.

[41] IAH demonstration years are not the same as calendar years (see Exhibit A.1).

- We assigned IAH beneficiaries whose practice TIN did not participate in an ACO at the beginning of the demonstration year to an ACO if they were attributed to an ACO in the beneficiary-level data in the MDM at any point during that study year.
- We assigned comparison beneficiaries to an ACO if they were attributed to an ACO in the beneficiary-level data in the MDM at any point during the study year.

7.2.3. ACO participation by IAH practices during the demonstration

Although IAH practices were not excluded from participating in ACOs before Year 3, CMS explicitly told practices that they could participate starting in Year 3 (which began in June 2014 for most practices). As a result, several IAH practices joined ACOs in Years 3, 4, and 5 of the IAH demonstration. Exhibit A.20 shows ACO participation for each IAH site at the beginning of each demonstration year according to information provided by IAH practices. No sites took part in an ACO before 2012.

The percentage of IAH beneficiaries assigned to an ACO for the purpose of our analysis increased substantially from Year 3 (19.8 percent) to Year 7 (82.1 percent) (Exhibit A.21). Among comparison beneficiaries, participation in an ACO increased steadily over time to 48.8 percent in Year 6, then declined to 40.6 by Year 7.

Exhibit A.20. IAH sites' ACO participation at the start of each IAH demonstration year

Demonstration practice location	Year 1 June 2012–May 2013	Year 2 June 2013–May 2014	Year 3 June 2014–May 2015	Year 4 Oct 2015–Sept 2016	Year 5 Oct 2016–Sept 2017	Year 6 Jan 2019–Dec 2019	Year 7 Jan 2020–Dec 2020
Practices that participated in Years 1 to 7							
Brooklyn, New York				✓	✓		
Dallas, Texas[a]				✓	✓	✓	✓
Flint, Michigan[a]	✓	✓	✓	✓	✓	✓	✓
Jacksonville, Florida[a]				✓	✓	✓	✓
Lansing, Michigan[a]				✓	✓	✓	✓
Long Island, New York					✓	✓	✓
Milwaukee, Wisconsin[a]				✓	✓	✓	✓
Philadelphia, Pennsylvania[b]							
Portland, Oregon							
Richmond, Virginia[b]							
Washington, DC[b]				✓	✓	✓	
Wilmington, Delaware				✓	✓	✓	✓
Practices that left the demonstration after Year 5[c]							
Austin, Texas						n.a.	n.a.
Cleveland, Ohio				✓	✓	n.a.	n.a.
Practices that left the demonstration after Year 6c							
Boston, Massachusetts				✓	✓	✓	n.a.
Durham, North Carolina							n.a.

Source: CMS's MDM database and information provided by IAH sites to Mathematica, supplemented with records from conversations with CMS and its implementation contractor and web searches.

Note: We do not report data for the first or second pre-demonstration years because none of the IAH practices participated in an ACO in those years (June 2010 to May 2011 and June 2011 to May 2012, respectively). As shown in Exhibit A.21, however, some IAH beneficiaries were assigned to an ACO in the second pre-demonstration year through attribution in the MDM.

[a] These practices participated as part of the Visiting Physicians Association.

[b] These practices participated as consortia and started Years 1 to 3 on September 1.

[c] Year 7 analyses included only the 10 practices that participated in that year. Similarly, Year 6 analyses included only the 12 practices that participated in that year. Analyses for Years 1 to 5 included the 14 practices that participated in all five years.

ACO = accountable care organization; MDM = Master Data Management system; n.a. = not applicable.

Exhibit A.21. Number and percentage of IAH and comparison beneficiaries assigned to an ACO, by evaluation year

	Second pre-demonstration year: June 2011–May 2012[a]	Year 1 June 2012–May 2013[a]	Year 2 June 2013–May 2014[a]	Year 3 June 2014–May 2015[a]	Year 4 Oct 2015–Sept 2016	Year 5 Oct 2016–Sept 2017	Year 6 Jan 2019–Dec 2019[b]	Year 7 Jan 2020–Dec 2020[b]
Number of IAH beneficiaries	107	1,773	1,392	1,498	7,001	7,445	5,509	4,153
Percentage of IAH beneficiaries	1.5	21.6	19.2	19.8	73.7	74.8	68.9	82.1
Number of comparison beneficiaries	1,246	5,412	8,431	10,685	15,846	19,670	15,872	8,468
Percentage of comparison beneficiaries	3.9	16.0	26.1	34.2	41.3	47.5	48.8	40.6

Source: CMS's MDM database and information provided by IAH sites to Mathematica, supplemented with records from conversations with CMS and its implementation contractor and web searches.

Note: We do not report data for the first pre-demonstration year (June 2010 to May 2011) because no IAH or comparison beneficiaries were assigned to an ACO that year.

[a] For the Richmond-based consortium, Years 1 to 3 and the two pre-demonstration years began in September rather than June.

[b] Year 7 analyses included only the 10 practices that participated in that year. Similarly, Year 6 analyses included only the 12 practices that participated in that year. Analyses for Years 1 to 5 included the 14 practices that participated in all five years.

ACO = accountable care organization; MDM = Master Data Management system.

7.2.4. Accounting for ACO participation in regression analysis

To assess whether ACO participation affected the impact of IAH on Medicare expenditures, we repeated our difference-in-differences regression controlling for a binary measure of ACO participation in each year, as we defined above. This enabled us to see whether the estimated effect of IAH differed depending on whether we controlled for ACO participation.

There are limitations to this analysis, largely related to potential measurement error associated with the measure of ACO participation we constructed. First, because our measure of participation in an ACO did not look at claims for individual beneficiaries, some IAH enrollees in the evaluation sample could have been misclassified as non-participants in an ACO. We assigned ACO participation to IAH beneficiaries based on attribution in the MDM as well as participation in an ACO at the practice level. IAH beneficiaries who received care from an IAH practice that was not affiliated with an ACO and who were enrolled in IAH in the MDM would not be considered ACO

participants for our analyses (except in rare cases in which an IAH enrollee was attributed to an ACO program in the MDM).

Second, we assigned IAH beneficiaries to an ACO based on their practice's ACO participation status at the start of the demonstration year. This approach could classify some beneficiaries as not in an ACO even though their IAH practice took part in an ACO for multiple months that year. For example, four IAH practices joined an ACO on January 1, 2015, partway through Year 3 of the IAH demonstration (June 2014 to May 2015). We classified beneficiaries associated with these practices as not being in an ACO for Year 3 because these practices were not affiliated with the ACO at the start of Year 3. If we measured these four practices as being part of an ACO in Year 3, we would have misclassified June to December 2014 as ACO months. Regardless of this distinction, if joining an ACO led IAH practices to experience substantial unmeasured changes in the patient population or care delivery that affected Medicare expenditures, such changes probably did not occur immediately after joining an ACO.

8. Qualitative methods and data

To understand why and how the incentive payments might (or might not) have affected outcomes, we had to understand how IAH practices' provision of home-based primary care changed after the IAH demonstration began and throughout the demonstration. Identifying the potential effect of changes by IAH practices also required understanding how the IAH participating practices provided home-based primary care before the IAH demonstration. Understanding the care delivery model enabled us to assess whether changes made by the participating practices appeared to be designed to reduce Medicare expenditures and improve health outcomes. When reporting information about the IAH practices in the report and appendices, we relied on qualitative data gathered from demonstration sites during demonstration Years 1 to 3, Year 5, and Year 6.

- We conducted the most recent interviews from November 2019 to February 2020 with 30 respondents. We interviewed one clinician and one practice administrator from each participating practice, except the one in Durham, North Carolina. In addition, we interviewed one corporate leader from the parent company of the VPA to collect additional insight on centralized supports for VPA practices. The Durham practice withdrew from the demonstration in fall 2019, so we conducted an exit interview with the practice administrator in December 2019. As part of this data collection effort, we used a semistructured interview protocol to collect information on changes practices made between our last round of interviews in April 2017 (during Year 5) and fall 2019 (during Year 6), including their efforts associated with meeting performance requirements for IAH quality measures. We also asked about motivations for any changes and factors that affected implementation of those changes.
- During the previous round of telephone interviews (conducted in April 2017), we interviewed 25 clinical and administrative staff at 15 IAH practices and the VPA corporate office in Troy, Michigan. We asked respondents about changes their practices had made during the demonstration to reduce hospital admissions and readmissions, reduce avoidable ED use, coordinate care, ensure round-the-clock access to care, follow up with patients and reconcile medications within 48 hours after discharge from the hospital or ED, and document patients' preferences. We also asked about motivation for making changes, clinician and staff reactions to changes, and factors that affected implementation of those changes.
- During telephone interviews conducted in January and February 2017, we collected information about IAH practices' structural characteristics and how they deliver care.
- During telephone interviews in late 2016 and early 2017, we interviewed 48 care partners across 13 of the 14 IAH practices that participated in Year 5, including at

least one home health agency that worked with each IAH practice.[42] Care partners are organizations external to the IAH practice's care team with which the practice has an established working relationship to coordinate care for patients. We asked respondents questions that enabled us to assess how care partner organizations perceive IAH practices on three key dimensions of home-based primary care: care coordination, accessibility, and continuity.

- During visits to demonstration sites from April 2015 to October 2015, we interviewed the sites' IAH team members and administrative staff involved in implementing the IAH demonstration. During this round of site visits, we focused on documenting changes in how the practices delivered care, the barriers to and facilitators of meeting the requirements of the demonstration, and how sites planned to sustain the home-based primary care model.
- Finally, we provide information gathered during earlier rounds of site visits: February to May 2013 (visits during Year 1) and February to July 2014 (visits during Year 2). During these earlier site visits, we focused on documenting how the practices delivered care, including changes from the year before the demonstration to Year 1 and changes from Year 1 to Year 2. During this period, we collected information on barriers to and facilitators of meeting the requirements of the demonstration and on how sites used information technologies, such as electronic health records and health information exchange to support their work.

For all interviews, we coded the data using a template that reflected the various requirements of the IAH demonstration (for example, providing patients with 24-hour access to the care team and working to reduce ED visits). The coding template also captured aspects of the five domains identified by the Consolidated Framework for Implementation Research (Damschroder et al. 2009) as playing an important role in implementation success: (1) the inner setting (internal attributes) of the practice sites, including structural and cultural characteristics affecting capacity for change; (2) the external environment (such as the availability of clinicians in the IAH practice's local market); (3) the characteristics of the IAH demonstration itself; (4) the characteristics of the people involved in implementing the model; and (5) the processes used to implement the model. We used qualitative analysis software (NVivo) to sort data using this coding template. We analyzed the sorted data to identify key barriers to and facilitators of implementation of the IAH demonstration in each participating site and identified common themes across sites.

[42] For more information about interviews with care partners, see the evaluation report covering Years 1 to 4 of the IAH demonstration.

In our analysis of qualitative data, we described what happened during the demonstration. We did not have a comparison group of primary care practices, so we could not be certain whether changes in practices' operations or structure occurred because of the demonstration. In addition, because we did not visit sites until after the demonstration began, data on practices' operations and structure before the demonstration was limited to what interviewees told us was different in Year 1 relative to before the demonstration.

9. Understanding differences between evaluation results and incentive payments

9.1. Approach for calculating incentive payments

Practices participating in the IAH demonstration are eligible to receive annual incentive payments. Specifically, practices can earn incentive payments if their patients' Medicare expenditures are below the practice's target expenditures and the practice meets performance thresholds for a set of quality measures. Calculation of incentive payments is based on (1) comparison of costs incurred (that is, actual expenditures of IAH-applicable beneficiaries during the performance year) with the target expenditures and (2) performance on payment-related quality measures. Actual expenditures of IAH-applicable beneficiaries are determined from Medicare FFS claims data. Target expenditures represent the expected expenditures based on Medicare FFS beneficiaries not exposed to the demonstration. Performance on payment-related quality measures is determined from information IAH practices report (for three of the six measures) and Medicare FFS claims data (for the remaining three measures).[43] IAH-applicable beneficiaries are those who are enrolled in the demonstration; for more information about IAH enrollees, see Chapter 3 of this appendix.

Although procedures for measuring performance on payment-related quality measures have remained unchanged throughout the IAH demonstration, calculations of actual and target expenditures have changed multiple times. Before beginning the demonstration, "CMS developed a risk-based actuarial methodology (the 'original actuarial methodology') for calculating incentive payments. In response to questions raised by participating IAH practices in early performance years regarding the risk scores used in the demonstration, CMS explored a different approach to the original actuarial method and developed a second methodology (the 'regression-based methodology'), which was later revised (the 'revised regression-based methodology')" (CMS 2022). For Year 5, calculations for 10 practices used the revised regression-based methodology, and calculations for 4 practices used the original actuarial methodology.[44]

For Years 6 and 7, calculations for all practices used the revised actuarial methodology, which generated "practice-specific PBPM target expenditures based on historical Medicare FFS per capita expenditures for the Medicare FFS population in the same counties as IAH-applicable beneficiaries. The per capita expenditures are adjusted to reflect the average CMS-HCC risk score, the average frailty score (used in the Program of All-inclusive Care for the Elderly), and a utilization factor of the IAH-

[43] See Exhibit 4.1 for a list of practice-reported and claims-based measures.

[44] More information on these methodologies is in the 2018 IAH Report to Congress, available at https://innovation.cms.gov/Files/reports/iah-rtc.pdf.

applicable population in each practice" (CMS 2022). The utilization factor is an add-on to the base risk score that "reflects the level of risk that was not captured by the CMS-HCC model for beneficiaries with a hospital admission and rehabilitation services use in the 12 months prior to their enrollment date in the performance year" (CMS 2022). New enrollees of IAH providers "receive a prospective CMS-HCC risk score, frailty factor, and utilization factor. The risk score and frailty factor for continuing enrollees are updated in future performance years only for changes in demographics (age and Medicaid status). The utilization factor is applied in future performance years only if a continuing enrollee had a hospitalization and post-acute care in the 12 months prior to the performance year" (CMS 2021b). The utilization factor led to an increase in target expenditures; all else equal, higher target expenditures result in a larger incentive payment. This implicitly assumes that IAH enrollees who were hospitalized and had post-acute care in one year are at equal risk for these outcomes in the following year. However, if some IAH enrollees tended to have a lower risk of hospitalization and post-acute care in the year after they had such utilization—in other words, if their expenditures tended to regress to the lower level of expenditures they incurred prior to the year that included a hospitalization and post-acute care—then the utilization factor may overestimate target expenditures. Finally, "the adjusted per capita expenditures are trended to the performance year by the increase in total per capita Medicare FFS expenditures, as estimated by CMS's Office of the Actuary" (CMS 2022). Trending expenditures forward to the performance year is necessary because the county-level data used to set target expenditures are lagged by about 15 months. For example, for Year 6, CMS used calendar year 2018 expenditures, which were published in April 2020. Trend factors come from the most recently available Medicare Trustees Report and are updated annually (CMS 2021b).

Equation (7) shows a simplified version of how CMS calculated the spending target for each beneficiary enrolled in the demonstration in Year 7. Equation (8) shows how CMS calculated the savings or loss PBPM for each practice.

(7) $$target\ \ expenditures = average\ \ monthly\ FFS\ county\ cost * cost\ trend * \left(beneficiary\ risk\ score + \ frailty\ \ factor + utilization\ \ factor\right)$$

(8) $$savings / loss\ PBPM = average\ \ target\ \ expenditures\ PBPM - average\ \ actual\ \ expenditures\ PBPM$$

The maximum incentive payment for each practice in Year 7 depended on the following factors (CMS 2021c):

- Savings or loss PBPM.
- Whether the estimated savings was statistically significant. "A one-sided confidence interval is constructed around each practice's actual expenditures, for

use in determining whether savings are statistically significant. [Target expenditures] are compared to the upper bound of the confidence intervals (80th and 85th) to determine if any observed savings are likely to be actual, rather than due to random variation" (CMS 2021b).

- The number of months of enrollment in the demonstration by IAH-applicable beneficiaries.
- Whether the maximum incentive payment exceeded the 5 percent savings kept by CMS, which was calculated as 5 percent multiplied by target expenditures and total beneficiary months.

If a practice had statistically significant savings that exceeded the 5 percent of savings that CMS retained, then the final incentive payment depended on the following:

- Whether the target expenditures were greater than the upper bound of the 85th confidence interval (rather than the 80th). If its target expenditures were greater than the upper bound of the 85th confidence interval, the practice earned 90 percent of the maximum payment. Otherwise, its target expenditures were greater than the upper bound of the 80th confidence interval, and the practice earned 50 percent of its target expenditures.
- The practice's performance on the six payment-related quality measures. If a practice met the performance thresholds for all six quality measures tied to payment, then it earned 100 percent of the maximum incentive payment. If a practice achieved the threshold for five, four, or three quality measures, it earned, respectively, 83 percent, 67 percent, or 50 percent of the maximum payment. The only requirement is that a practice must meet the performance threshold for the quality measures—payment does not vary by how much the practice exceeded the threshold.

As an example, consider a practice that had a $1 million maximum incentive payment after subtracting the 5 percent of savings that CMS retains and both of the following: (1) its target expenditures were greater than the upper bound of the 85th confidence interval and (2) it achieved the threshold for three quality measures. For this practice, the actual incentive payment would be $1 million multiplied by 80 percent (for the 85th confidence interval) multiplied by 50 percent (for achieving the threshold for three quality measures), for a final payment of $400,000.

9.2. Differences between evaluation results and incentive payments

9.2.1. Purposes of the evaluation and incentive payment calculation

The evaluation and the incentive payment calculation serve different purposes. The evaluation needs to estimate the effect of the demonstration after accounting for how factors other than the demonstration that changed during the performance period affected expenditures. Conducting a rigorous evaluation requires using the

same data and approach to identify IAH and comparison groups and advanced statistical methods to reduce the risk of bias in the study results. Specifically, it requires using a comparison group of beneficiaries that are similar to IAH beneficiaries but did not receive home-based primary care and—as we describe in the following section—examining changes in expenditures for the comparison group relative to changes for patients of IAH practices over the same time period. By contrast, the incentive payment calculation needs to offer a timely way to measure a target spending level to reward participation, subject to other requirements such as quality measure performance. As described in the previous section of this chapter, this is done by trending historical expenditures for Medicare FFS beneficiaries and adjusting those expenditures to reflect IAH beneficiaries' health status and past utilization.

9.2.2. Reasons why the evaluation and incentive payment calculation results differ each year

In all years of the demonstration, large differences have existed between the total incentive payments paid to IAH practices and the estimated aggregate effect of IAH calculated by the evaluation (see Exhibit 2.4 of the report for more information). In Year 7, the estimated aggregated effect was a reduction of $22.6 million, and total incentive payments were $18.5 million (Exhibit A.22). The evaluation and incentive payment calculation results differ each year because of (1) differences in the sample of beneficiaries included and (2) differences in the methods used. As we explain in this section, our use of a difference-in-differences approach accounts for much of the discrepancy between evaluation findings and incentive payment calculation results.

To evaluate the effect of IAH, Mathematica used a quasi-experimental difference-in-differences design. Under this design, we estimated effects as the change in outcomes for IAH-eligible beneficiaries receiving care from IAH practices before and after the start of the demonstration relative to the change during the same period for a matched comparison group. Equation (9) shows a simplified version of how the evaluation calculated the effect of IAH on expenditures in Year 7. The terms IAH_{y7} and C_{y7} reflect regression-adjusted PBPM expenditures in Year 7 for IAH and comparison beneficiaries, respectively. The terms IAH_{pd} and C_{pd} reflect regression-adjusted PBPM expenditures in the pre-demonstration year 2011–2012 for IAH and comparison beneficiaries, respectively; this is the year we used as the baseline for calculating the difference-in-differences estimate.

(9) $$effect\ on\ expenditures = (IAH_{y7} - C_{y7}) - (IAH_{pd} - C_{pd})$$

In Year 7, regression-adjusted PBPM expenditures were $5,002 for IAH beneficiaries and $5,862 for the comparison group, for a difference of –$860 PBPM (Exhibit A.22).

This difference of –$860 PBPM is conceptually similar to the difference between actual and target expenditures used in the incentive payment calculation, where the difference between actual expenditures for IAH-applicable beneficiaries and target expenditures was –$693 PBPM for the average beneficiary in Year 7. In other words, $IAH_{y7} - C_{y7}$ in Equation (9) is conceptually similar to Equation (8). The difference between the IAH-comparison group difference in Year 7 calculated by the evaluation (–$860 PBPM) and that calculated for incentive payments (–$693 PBPM) was –$167 PBPM, or 19 percent. It is likely that each of the differences in sample (identification of IAH beneficiaries and a comparison or target group)[45] and methods (calculation of expenditures) contributed at least somewhat to the -$167 PBPM difference.

Unlike the incentive payment calculation, the evaluation seeks to estimate the *change* in the difference between IAH and comparison beneficiaries relative to before the demonstration began. Before the demonstration began, beneficiaries who met IAH eligibility criteria and were treated by IAH practices in the year before the demonstration had lower expenditures than a matched comparison group of beneficiaries who had similar observable characteristics and health status but did not receive home-based primary care. One year before the demonstration, regression-adjusted PBPM expenditures for the 10 practices that participated in Year 7 were $4,965 for beneficiaries eligible for IAH and $5,366 for the comparison group, for a difference of –$401; the difference between –$860 (from Year 7) and –$401 is –$459, which is the evaluation's estimated effect of IAH in Year 7.

Why did the evaluation use the change in the difference between IAH and comparison beneficiaries rather than the IAH-comparison difference in Year 7 only? First, interpreting the difference in mean expenditures for IAH and comparison beneficiaries in a single demonstration year as the effect of the demonstration poses a substantial risk of bias by not accounting for any potential unobserved factors that affected expenditures unrelated to the payment incentive and delivery of home-based primary care. For example, it is possible that unobserved differences between IAH beneficiaries and comparison beneficiaries influenced both the decision to start home-based primary care at an IAH practice and health care spending patterns. These could include differences in health status and environment that were not captured in Medicare claims or administrative data, as well as differences in attitudes and preferences about health care. Using a difference-in-differences approach avoids this type of bias as long as unobserved differences between IAH beneficiaries and comparison beneficiaries, and the influence of those unobservable variables on outcomes, are consistent over time. Also, by using data from both the baseline and

[45] For more information about differences in the IAH samples used for the evaluation and the incentive payment calculation, see Chapter 3 of this appendix.

Year 7 for both IAH and comparison beneficiaries, this approach accounts for underlying trends that affect Medicare expenditures the same way for both groups.

Second, while the demonstration has two components—a payment incentive and the delivery of home-based primary care—only the payment incentive has the potential to generate expenditure reductions for Medicare that *would not have been achieved without the IAH demonstration*. This is because payment for home-based primary care was in effect before the demonstration. Prior to the COVID-19 pandemic—that is, for IAH Years 1 to 6—using the change in the difference between IAH and comparison beneficiaries allowed us to measure the expenditure reduction (if any) for Medicare that was achieved solely because of the payment incentive available during the IAH demonstration. But this interpretation does not hold in the first year of the pandemic, primarily because the relative effectiveness care for IAH and comparison beneficiaries may have changed (see Chapter 4 of the report for more details and examples) as a result of disruption in health care and society more generally during the pandemic. In Year 7, the estimated effects reflect both of the following:

- Any effects of changes in care delivery by IAH practices because of the IAH payment incentive—that is, potential Medicare spending reductions that *would not have* been achieved without the IAH demonstration (the focus of the evaluation in Years 1 to 6).
- Any effects of changes in the relative effectiveness of care for IAH and comparison beneficiaries during the first year of the pandemic—effects which *would have* been achieved during the pandemic without the IAH demonstration, since the IAH practices offered home-based primary care before the demonstration began.

9.2.3. Reasons why the difference in the evaluation and incentive payment calculation results changed drastically between Years 6 and 7

The difference of –$860 PBPM between the IAH and comparison groups used for the evaluation for Year 7 was 19 percent larger than the difference between IAH practices' spending and their spending targets used for the incentive payment calculation (-$693 PBPM) (Exhibit A.22). Similarly, the estimated reduction in aggregate spending calculated by the evaluation was about 18 percent larger than total incentive payments ($22.6 million and $18.5 million, respectively). This amount differed sharply from Year 6. In Year 6, the estimated reduction in aggregate spending calculated by the evaluation was about 250 percent *smaller* than total incentive payments ($3.2 million and $11.1 million, respectively), despite the fact that the difference between the IAH and comparison groups calculated by the evaluation was only 28 percent smaller than that calculated for the incentive payment calculation.

The stark change between the differences in aggregate spending calculated by the evaluation and total incentive payments—an 18 percent difference in Year 7 and a 250 percent difference in Year 6—occurred primarily because the estimated effect in Year 7 calculated by the evaluation (-$459) was more than 10 times larger than in Year 6 (-$41). By contrast, the estimated spending difference used to calculate incentive payments was only 31 percent larger in Year 7 (-$693) than in Year 6 (-$529). As discussed earlier in this section, estimated effects in Year 7 reflect any effects of changes in the relative effectiveness of care for IAH and comparison beneficiaries during the first year of the pandemic—a feature which did not apply to Years 1 to 6 or to the year before the demonstration.

Exhibit A.22. Comparison of evaluation and incentive payment calculation results in Years 6 and 7

	Year 7		Year 6	
	Evaluation[a]	Incentive payment calculation[b]	Evaluation[a]	Incentive payment calculation[b]
IAH beneficiaries' spending, PBPM	$5,002	$3,159	$4,725	$3,314
Comparison beneficiaries' spending (evaluation) or spending target (incentive payment calculation), PBPM	$5,862	$3,852	$5,137	$3,843
Absolute difference between IAH and comparison (evaluation) or IAH and spending target (incentive payment calculation), PBPM	-$860	-$693	-$412	-$529
Relative difference between evaluation and incentive payment calculation	19%		-28%	
Difference between IAH and comparison beneficiaries in the year before the demonstration, PBPM	-$401	n.a.	-$371	n.a.
Estimated effect of IAH used to calculate aggregate effects for the evaluation, PBPM	-$459	n.a.	-$41	n.a.
Estimated spending difference used to calculate incentive payments	n.a.	-$693	n.a.	-$529
Number of beneficiary-months used to calculate estimated aggregate spending reduction (evaluation) or incentive payments given to IAH practices (incentive payment calculation)[c]	49,301	61,128	77,224	60,977
Estimated aggregate spending reduction (evaluation) or incentive payments given to IAH practices (incentive payment calculation)	-$22.6 million[d]	$18.5 million	-$3.2 million[d]	$11.1 million
Relative difference between evaluation and incentive payment calculation	-18%		247%	

Sources: Mathematica's analysis of data from the IAH implementation contractor and Medicare claims and enrollment data from the Chronic Conditions Warehouse. Results for the incentive payment calculation are provided by CMS at https://innovation.cms.gov/initiatives/independence-at-home.

Exhibit A.22 (*continued*)

Note: Differences between Years 5, 6, and 7 represent the change in participating sites as well as any differences before and during the COVID-19 pandemic in the effects of the IAH payment incentive and home-based primary care. Results cannot be generalized to other years or home-based primary care providers.

[a] For the evaluation, Year 7 reflects the 10 practices that participated in Year 7, and Year 6 reflects the 12 practices that participated in Year 6.

[b] For the incentive payment calculation, Year 7 reflects the 10 practices that participated in Year 7, and Year 6 reflects the 11 practices that completed Year 6.

[c] See Chapter 3 of this appendix for details about the differences between the IAH beneficiaries used for the evaluation and IAH enrollees used for the incentive payment calculation. One very large practice participated in most, but not all, of Year 6. This practice was included in the evaluation of Year 6 but not the incentive payment calculation. This is the primary reason why the number of beneficiary months for the evaluation decreased sharply from Year 6 to Year 7, whereas the number of beneficiary months for the incentive payment calculation was similar in those two years. See Chapter 3 of this appendix for discussion about other factors that led to us identifying fewer IAH-eligible beneficiaries for the evaluation in Year 7.

[d] Estimated aggregate spending reductions calculated by the evaluation reflect the beneficiary-level estimated effect of IAH multiplied by the number of IAH beneficiary-months in each year. Incentive payment calculations reflect the difference between spending for IAH enrollees and the spending target, number of beneficiary-months, and other features discussed earlier in this chapter (such as CMS keeping the first 5 percent of savings and the practice's performance on the six payment-related quality measures).

n.a. = not applicable; PBPM = per beneficiary per month.

Appendix B: Effects of IAH – Supplementary Exhibits

Exhibit B.1. Baseline unadjusted means of outcomes among all IAH beneficiaries

Outcome name	Practices that participated in Year 6	Practices that participated in Year 7
PBPM Medicare spending		
Total	$4,186	$4,287
Inpatient	$1,627	$1,690
Skilled nursing facility	$621	$609
Home health (Parts A and B)	$717	$738
Hospice	$162	$162
Outpatient	$247	$246
Clinician/supplier	$675	$697
Durable medical equipment	$136	$145
Service use outcomes		
Number of hospital admissions per 1,000 beneficiaries per year[a]	1,727	1,767
Number of potentially avoidable hospital admissions per 1,000 beneficiaries per year[b]	439	450
Probability (as a percentage) of having a qualifying hospital discharge and unplanned readmission within 30 days of discharge	15.91	19.69
Number of outpatient ED visits per 1,000 beneficiaries per year[c]	1,440	1,413
Number of hospital admissions preceded by an ED visit per 1,000 beneficiaries per year	1,422	1,447
Number of potentially avoidable outpatient ED visits per beneficiary per year[d]	185	187
Health outcomes		
Probability (as a percentage) of dying in the demonstration year	14.90	16.78
Probability (as a percentage) of entering institutional long-term care in the demonstration year	12.46	12.56

Source: Mathematica's analysis of data from the IAH implementation contractor and Medicare claims and enrollment data from the Chronic Conditions Warehouse.

Notes: We calculated the baseline means of outcomes using the sample of IAH-eligible beneficiaries in the year before IAH started. We report the baseline means for two samples: the 12 practices that participated in Year 6 and the 10 practices that participated in Year 7.

[a] The number of hospital admissions includes observation stays.

[b] The number of potentially avoidable hospital admissions includes observation stays. A potentially avoidable hospital admission is one in which appropriate primary and specialty care may prevent or reduce the need for a hospital admission.

[c] The number of outpatient ED visits measures all those not resulting in a hospital admission, including those resulting in an observation stay.

[d] A potentially avoidable outpatient ED visit is one in which appropriate primary and specialty care may prevent or reduce the need for such visits. The measure excluded ED visits that led to a hospital admission because there was no diagnosis from such a visit in a claim record when the ED visit led to an admission.

Exhibit B.1 (*continued*)

ED = emergency department; n.a. = not applicable; PBPM = per beneficiary per month.

Exhibit B.2a. Estimated effect of IAH on total Medicare spending PBPM in Years 1 to 7, practices that participated in Year 7

	IAH	Comparison	Difference (IAH - comparison)	Difference-in-differences estimated effect	Percentage effect[a]
Year 7	$5,002	$5,862	-$860	-$459** ($206)	-10.7
Year 6	$5,070	$5,618	-$548	-$148 ($150)	-3.4
Year 5	$4,961	$5,485	-$524	-$123 ($179)	-2.9
Year 4	$4,767	$5,157	-$390	$11 ($202)	0.2
Year 3	$4,894	$5,183	-$289	$112 ($129)	2.6
Year 2	$5,035	$5,233	-$198	$203 ($127)	4.7
Year 1	$4,967	$5,232	-$265	$136 ($103)	3.2
One year pre-IAH[b]	$4,965	$5,366	-$401	-	-
Two years pre-IAH	$ 5,123	$ 5,567	-$444	-$44 ($72)	-1.0

Source: Mathematica's analysis of data from the IAH implementation contractor and Medicare claims and enrollment data from the Chronic Conditions Warehouse.

Notes: Total unweighted number of observations across all years is 259,609. Standard errors are in parentheses. We computed coefficients and standard errors by using the weighted sample size, which considers the matching and eligibility weights. The table reports the regression-adjusted means of the IAH and matched comparison groups in each year. We obtained regression-adjusted means by applying the estimated regression coefficients to the covariates of IAH beneficiaries in the latest demonstration year. Because of rounding, a difference-in-differences estimate displayed as zero may be shown alongside a percentage effect that exceeds zero. Estimated effects in Year 7 are interpreted as effects of IAH during the first year of the COVID-19 pandemic.

*/**/*** The difference is statistically significant at the 0.10/0.05/0.01 level.

[a] We used the unadjusted IAH group mean in the year before the demonstration to calculate the percentage effect for each demonstration year. Exhibit B.1 reports the baseline unadjusted IAH group mean for all outcomes.

[b] The difference-in-differences estimate for the period before the demonstration is zero (with no standard error) in all regressions because we calculated that estimate for each year as the difference in means between IAH and comparison beneficiaries in that year minus the difference in the year before the demonstration.

PBPM = per beneficiary per month.

Exhibit B.2b. Estimated effect of the IAH payment incentive on total Medicare spending PBPM in Years 1 to 6, practices that participated in Year 6

	IAH	Comparison	Difference (IAH - comparison)	Difference-in-differences estimated effect	Percentage effect[a]
Year 6	$4,725	$5,137	-$412	-$41 ($126)	-1.0
Year 5	$4,530	$5,014	-$483	-$113 ($157)	-2.7
Year 4	$4,326	$4,715	-$389	-$19 ($179)	-0.4
Year 3	$4,473	$4,791	-$318	$52 ($117)	1.3
Year 2	$4,610	$4,830	-$220	$150 ($122)	3.6
Year 1	$4,587	$4,857	-$270	$101 ($92)	2.4
One year pre-IAH[b]	$4,626	$4,996	-$371	-	-
Two years pre-IAH	$4,811	$5,209	-$398	-$27 ($64)	-0.6

Source: Mathematica's analysis of data from the IAH implementation contractor and Medicare claims and enrollment data from the Chronic Conditions Warehouse.

Notes: Total unweighted number of observations across all years is 290,514. Standard errors are in parentheses. We computed coefficients and standard errors by using the weighted sample size, which considers the matching and eligibility weights. The table reports the regression-adjusted means of the IAH and matched comparison groups in each year. We obtained regression-adjusted means by applying the estimated regression coefficients to the covariates of IAH beneficiaries in the latest demonstration year. Because of rounding, a difference-in-differences estimate displayed as zero may be shown alongside a percentage effect that exceeds zero.

*/**/*** The difference is statistically significant at the 0.10/0.05/0.01 level.

[a] We used the unadjusted IAH group mean in the year before the demonstration to calculate the percentage effect for each demonstration year. Exhibit B.1 reports the baseline unadjusted IAH group mean for all outcomes.

[b] The difference-in-differences estimate for the period before the demonstration is zero (with no standard error) in all regressions because we calculated that estimate for each year as the difference in means between IAH and comparison beneficiaries in that year minus the difference in the year before the demonstration.

PBPM = per beneficiary per month.

Exhibit B.3. Estimated effect of IAH on total Medicare spending PBPM under Bayesian model in Years 1 to 7, practices that participated in Year 7

	Difference-in-differences estimated effect	Percentage effect[a]	Probability of reducing spending at least $100 PBPM	Probability of reducing spending by at least $200 PBPM
Year 7	-$446 (-$707, -$187)	-10.4	98.4%	94.0%
Year 6	$3 (-$254, $259)	0.1	25.3%	9.6%
Year 5	-$119 (-$375, $133)	-2.8	55.1%	30.2%
Year 4	-$28 (-$284, $226)	-0.7	31.9%	13.3%
Year 3	$36 (-$219, $290)	0.8	19.2%	6.4%
Year 2	$172 (-$89, $426)	4.0	4.4%	1.0%
Year 1	$16 (-$238, $276)	0.4	22.7%	8.1%
One year pre-IAH[b]	-	-	-	-
Two years pre-IAH	-$71 (-$332, $186)	-1.7	42.6%	20.6%

Source: Mathematica's analysis of data from the IAH implementation contractor and Medicare claims and enrollment data from the Chronic Conditions Warehouse.

Notes: We report the 90 percent credible intervals in parentheses. We computed coefficients and standard errors by using the weighted sample size, which considers the matching and eligibility weights. The table reports the regression-adjusted means of the IAH and matched comparison groups in each year. We obtained regression-adjusted means by applying the estimated regression coefficients to the covariates of IAH beneficiaries in the latest demonstration year. Because of rounding, a difference-in-differences estimate displayed as zero may be shown alongside a percentage effect that exceeds zero. Estimated effects in Year 7 are interpreted as effects of IAH during the first year of the COVID-19 pandemic.

[a] We used the unadjusted IAH group mean in the year before the demonstration to calculate the percentage effect for each demonstration year. Exhibit B.1 reports the baseline unadjusted IAH group mean for all outcomes.

[b] The difference-in-differences estimate for the period before the demonstration is zero (with no credible interval) in all regressions because we calculated that estimate for each year as the difference in means between IAH and comparison beneficiaries in that year minus the difference in the year before the demonstration.

PBPM = per beneficiary per month.

Exhibit B.4a. Estimated effect of IAH on Medicare spending PBPM in Years 1 to 7, by service category, practices that participated in Year 7

Service type	IAH	Comparison	Difference (IAH - comparison)	Difference-in-differences estimated effect	Percentage effect[a]
Inpatient					
Year 7	$2,161	$2,791	-$630	-$245* ($138)	-14.5
Year 6	$2,147	$2,610	-$463	-$77 ($97)	-4.6
Year 5	$2,110	$2,560	-$451	-$65 ($126)	-3.9
Year 4	$2,038	$2,347	-$309	$76 ($129)	4.5
Year 3	$2,079	$2,354	-$275	$111 ($93)	6.6
Year 2	$2,124	$2,393	-$269	$116 ($79)	6.9
Year 1	$2,099	$2,360	-$261	$124* ($71)	7.3
One year pre-IAH[b]	$2,045	$2,431	-$385	-	
Two years pre-IAH	$2,116	$2,513	-$398	-$13 ($63)	-0.7
SNF					
Year 7	$663	$968	-$305	-$63 ($54)	-10.4
Year 6	$684	$895	-$211	$32 ($39)	5.2
Year 5	$723	$938	-$215	$27 ($54)	4.5
Year 4	$700	$935	-$236	$7 ($62)	1.1
Year 3	$717	$925	-$207	$35 ($35)	5.7
Year 2	$693	$895	-$201	$41 ($32)	6.7
Year 1	$679	$905	-$227	$16 ($25)	2.6
One year pre-IAH[b]	$681	$923	-$242	-	-
Two years pre-IAH	$727	$993	-$266	-$24 ($24)	-3.9
Home health (Parts A and B)					
Year 7	$742	$532	$210	$13 ($30)	1.8
Year 6	$694	$508	$186	-$10 ($40)	-1.4
Year 5	$724	$517	$207	$11 ($38)	1.4
Year 4	$679	$498	$181	-$15 ($38)	-2.1
Year 3	$691	$514	$178	-$19 ($38)	-2.6
Year 2	$812	$581	$231	$35 ($26)	4.7
Year 1	$757	$574	$183	-$14 ($20)	-1.8
One year pre-IAH[b]	$783	$586	$196	-	
Two years pre-IAH	$838	$638	$200	$4 ($12)	0.5

Exhibit B.4a (*continued*)

Service type	IAH	Comparison	Difference (IAH - comparison)	Difference-in-differences estimated effect	Percentage effect[a]
Hospice					
Year 7	$169	$122	$47	-$23 ($14)	-14.1
Year 6	$209	$112	$96	$26* ($14)	16.1
Year 5	$170	$105	$65	-$5 ($14)	-3.3
Year 4	$165	$106	$59	-$11 ($11)	-7.0
Year 3	$170	$98	$72	$2 ($17)	1.3
Year 2	$154	$82	$71	$1 ($14)	0.9
Year 1	$169	$102	$67	-$3 ($9)	-1.6
One year pre-IAH[b]	$172	$102	$70	-	-
Two years pre-IAH	$144	$95	$49	-$21** ($10)	-12.8
Outpatient					
Year 7	$324	$463	-$139	-$68*** ($11)	-27.7
Year 6	$354	$500	-$145	-$75*** ($13)	-30.3
Year 5	$333	$439	-$106	-$35*** ($10)	-14.4
Year 4	$313	$414	-$100	-$30*** ($11)	-12.1
Year 3	$322	$405	-$83	-$12 ($11)	-4.9
Year 2	$317	$388	-$70	$0 ($10)	0.1
Year 1	$295	$373	-$79	-$8 ($8)	-3.1
One year pre-IAH[b]	$301	$372	-$71	-	
Two years pre-IAH	$300	$358	-$58	$13 ($9)	5.2
Clinician/supplier					
Year 7	$798	$861	-$63	-$61 ($39)	-8.7
Year 6	$836	$868	-$32	-$30 ($39)	-4.3
Year 5	$787	$812	-$25	-$24 ($24)	-3.4
Year 4	$757	$753	$4	$6 ($35)	0.9
Year 3	$785	$770	$15	$17 ($26)	2.5
Year 2	$797	$776	$21	$23 ($24)	3.3
Year 1	$801	$775	$26	$28* ($16)	4.0
One year pre-IAH[b]	$801	$802	-$2	-	
Two years pre-IAH	$803	$806	-$3	-$1 ($12)	-0.1
Durable medical equipment					
Year 7	$146	$126	$20	-$13 ($16)	-9.0
Year 6	$145	$125	$20	-$13 ($9)	-9.2
Year 5	$115	$113	$2	-$31*** ($7)	-21.6
Year 4	$115	$104	$11	-$22*** ($8)	-15.4
Year 3	$129	$118	$11	-$22*** ($8)	-15.5

Exhibit B.4a (*continued*)

Service type	IAH	Comparison	Difference (IAH - comparison)	Difference-in-differences estimated effect	Percentage effect[a]
Year 2	$138	$117	$20	-$13* ($7)	-8.9
Year 1	$168	$143	$25	-$8 ($5)	-5.7
One year pre-IAH[b]	$184	$151	$33	-	-
Two years pre-IAH	$195	$164	$31	-$2 ($7)	-1.5

Source: Mathematica's analysis of data from the IAH implementation contractor and Medicare claims and enrollment data from the Chronic Conditions Warehouse.

Notes: Total unweighted number of observations across all years is 259,609. Standard errors are in parentheses. We computed coefficients and standard errors by using the weighted sample size, which considers the matching and eligibility weights. The table reports the regression-adjusted means of the IAH and matched comparison groups in each year. We obtained regression-adjusted means by applying the estimated regression coefficients to the covariates of IAH beneficiaries in the latest demonstration year. We calculated percentage effects using the unadjusted IAH group mean in the year before the demonstration. Because of rounding, a difference-in-differences estimate displayed as zero may be shown alongside a percentage effect that exceeds zero. Estimated effects in Year 7 are interpreted as effects of IAH during the first year of the COVID-19 pandemic.

*/**/*** The difference is statistically significant at the 0.10/0.05/0.01 level.

[a] We used the unadjusted IAH group mean in the year before the demonstration to calculate the percentage effect for each demonstration year. Exhibit B.1 reports the baseline unadjusted IAH group mean for all outcomes.

[b] The difference-in-differences estimate for the year before the demonstration is zero (with no standard error) in all regressions because we calculated that estimate for each year as the difference in means between IAH and comparison beneficiaries in that year minus the difference in the year before the demonstration.

PBPM = per beneficiary per month; SNF = skilled nursing facility.

Exhibit B.4b. Estimated effect of the IAH payment incentive on Medicare spending PBPM in Years 1 to 6, by service category, practices that participated in Year 6

Service type	IAH	Comparison	Difference (IAH - comparison)	Difference-in-differences estimated effect	Percentage effect[a]
Inpatient					
Year 6	$1,876	$2,313	-$436	-$66 ($90)	-4.1
Year 5	$1,815	$2,263	-$448	-$77 ($116)	-4.8
Year 4	$1,718	$2,067	-$349	$21 ($130)	1.3
Year 3	$1,805	$2,109	-$304	$67 ($88)	4.1
Year 2	$1,854	$2,157	-$303	$67 ($79)	4.1
Year 1	$1,864	$2,135	-$271	$99 ($65)	6.1
One year pre-IAH[b]	$1,839	$2,209	-$370	-	
Two years pre-IAH	$1,927	$2,303	-$376	-$6 ($54)	-0.3
SNF					
Year 6	$687	$882	-$196	$36 ($35)	5.8
Year 5	$713	$929	-$217	$15 ($52)	2.4
Year 4	$694	$927	-$233	-$1 ($56)	-0.2
Year 3	$725	$938	-$213	$19 ($32)	3.1
Year 2	$711	$902	-$191	$40 ($28)	6.5
Year 1	$691	$919	-$228	$4 ($22)	0.6
One year pre-IAH[b]	$706	$938	-$232	-	-
Two years pre-IAH	$755	$1,008	-$253	-$21 ($22)	-3.4
Home health (Parts A and B)					
Year 6	$662	$475	$187	-$9 ($38)	-1.2
Year 5	$689	$485	$204	$8 ($35)	1.2
Year 4	$647	$460	$187	-$8 ($36)	-1.2
Year 3	$649	$473	$176	-$20 ($34)	-2.7
Year 2	$762	$535	$227	$31 ($23)	4.4
Year 1	$717	$531	$186	-$10 ($18)	-1.3
One year pre-IAH[b]	$739	$544	$196	-	
Two years pre-IAH	$798	$595	$204	$8 ($12)	1.1
Hospice					
Year 6	$224	$130	$94	$30** ($14)	18.2
Year 5	$174	$117	$57	-$8 ($11)	-4.7
Year 4	$177	$119	$58	-$6 ($12)	-3.8
Year 3	$171	$107	$64	$0 ($14)	-0.1
Year 2	$156	$89	$67	$3 ($12)	1.6
Year 1	$171	$111	$60	-$4 ($9)	-2.4

Exhibit B.4b (*continued*)

Service type	IAH	Comparison	Difference (IAH - comparison)	Difference-in-differences estimated effect	Percentage effect[a]
One year pre-IAH[b]	$176	$112	$64	-	-
Two years pre-IAH	$150	$105	$45	-$19** ($9)	-11.7
Outpatient					
Year 6	$361	$457	-$96	-$31 ($26)	-12.6
Year 5	$324	$401	-$77	-$12 ($15)	-5.0
Year 4	$301	$380	-$79	-$15 ($15)	-6.0
Year 3	$300	$368	-$68	-$4 ($10)	-1.4
Year 2	$290	$352	-$62	$3 ($9)	1.1
Year 1	$268	$340	-$71	-$7 ($8)	-2.7
One year pre-IAH[b]	$274	$339	-$65	-	
Two years pre-IAH	$271	$323	-$51	$13 ($8)	5.4
Clinician/supplier					
Year 6	$803	$779	$24	$17 ($38)	2.6
Year 5	$727	$729	-$2	-$9 ($21)	-1.3
Year 4	$699	$679	$21	$14 ($30)	2.1
Year 3	$722	$699	$23	$16 ($23)	2.4
Year 2	$727	$702	$25	$19 ($21)	2.7
Year 1	$736	$705	$31	$24* ($14)	3.6
One year pre-IAH[b]	$737	$731	$6	-	
Two years pre-IAH	$745	$739	$6	$0 ($11)	0.0
Durable medical equipment					
Year 6	$112	$101	$11	-$19** ($8)	-13.6
Year 5	$89	$90	$0	-$30*** ($7)	-22.3
Year 4	$89	$83	$6	-$24*** ($8)	-17.4
Year 3	$101	$97	$4	-$26*** ($7)	-19.2
Year 2	$110	$93	$17	-$13** ($6)	-9.5
Year 1	$139	$116	$23	-$7 ($5)	-5.1
One year pre-IAH[b]	$154	$124	$30	-	-
Two years pre-IAH	$164	$136	$28	-$2 ($6)	-1.6

Source: Mathematica's analysis of data from the IAH implementation contractor and Medicare claims and enrollment data from the Chronic Conditions Warehouse.

Notes: Total unweighted number of observations across all years is 290,514. Standard errors are in parentheses. We computed coefficients and standard errors by using the weighted sample size, which considers the matching and eligibility weights. The table reports the regression-adjusted means of the IAH and matched comparison groups in each year. We obtained regression-adjusted means by applying the estimated regression coefficients to the covariates of IAH beneficiaries in the latest demonstration year. We calculated percentage effects using the unadjusted IAH group mean in the year before the demonstration. Because of

Exhibit B.4b (*continued*)

rounding, a difference-in-differences estimate displayed as zero may be shown alongside a percentage effect that exceeds zero.

*/**/*** The difference is statistically significant at the 0.10/0.05/0.01 level.

[a] We used the unadjusted IAH group mean in the year before the demonstration to calculate the percentage effect for each demonstration year. Exhibit B.1 reports the baseline unadjusted IAH group mean for all outcomes.

[b] The difference-in-differences estimate for the year before the demonstration is zero (with no standard error) in all regressions because we calculated that estimate for each year as the difference in means between IAH and comparison beneficiaries in that year minus the difference in the year before the demonstration.

PBPM = per beneficiary per month; SNF = skilled nursing facility.

Exhibit B.5a. Estimated effect of IAH on total Medicare spending PBPM in Years 1 to 7, controlling for COVID-19 diagnosis in Year 7

	IAH	Comparison	Difference (IAH - comparison)	Difference-in-differences estimated effect	Percentage effect[a]
Year 7	$5,002	$5,848	-$846	-$448** ($204)	-10.4
Year 6	$5,483	$6,030	-$547	-$149 ($150)	-3.5
Year 5	$5,374	$5,896	-$523	-$124 ($179)	-2.9
Year 4	$5,181	$5,569	-$388	$11 ($202)	0.2
Year 3	$5,306	$5,592	-$287	$112 ($129)	2.6
Year 2	$5,447	$5,642	-$196	$203 ($127)	4.7
Year 1	$5,377	$5,640	-$263	$136 ($104)	3.2
One year pre-IAH[a]	$5,377	$5,775	-$398	-	-
Two years pre-IAH	$5,532	$5,974	-$442	-$44 ($72)	-1.0

Source: Mathematica's analysis of data from the IAH implementation contractor and Medicare claims and enrollment data from the Chronic Conditions Warehouse.

Notes: Total unweighted number of observations across all years is 259,609. The models used to estimate effects in this table are identical to those used in Exhibit B.2a with the exception of a single binary indicator for being diagnosed with COVID-19 any time in Year 7 (2020). Standard errors are in parentheses. We computed coefficients and standard errors by using the weighted sample size, which considers the matching and eligibility weights. The table reports the regression-adjusted means of the IAH and matched comparison groups in each year. We obtained regression-adjusted means by applying the estimated regression coefficients to the covariates of IAH beneficiaries in the latest demonstration year. We calculated percentage effects using the unadjusted IAH group mean in the year before the demonstration. Because of rounding, a difference-in-differences estimate displayed as zero may be shown alongside a percentage effect that exceeds zero. Estimated effects in Year 7 are interpreted as effects of IAH during the first year of the COVID-19 pandemic.

*/**/*** The difference is statistically significant at the 0.10/0.05/0.01 level.

[a] The difference-in-differences estimate for the year before the demonstration is zero (with no standard error) in all regressions because we calculated that estimate for each year as the difference in means between IAH and comparison beneficiaries in that year minus the difference in the year before the demonstration.

PBPM = per beneficiary per month.

Exhibit B.5b. Estimated effect of IAH on total Medicare spending PBPM in Years 1 to 7, controlling for COVID-19 hospitalization in Year 7

	IAH	Comparison	Difference (IAH - comparison)	Difference-in-differences estimated effect	Percentage effect[a]
Year 7	$5,002	$5,869	-$867	-$468 ($209)	-10.9
Year 6	$5,359	$5,907	-$548	-$149 ($150)	-3.5
Year 5	$5,250	$5,773	-$523	-$124 ($179)	-2.9
Year 4	$5,057	$5,445	-$388	$11 ($202)	0.3
Year 3	$5,182	$5,469	-$287	$112 ($129)	2.6
Year 2	$5,323	$5,519	-$196	$203 ($127)	4.7
Year 1	$5,254	$5,518	-$263	$136 ($104)	3.2
One year pre-IAH[a]	$5,254	$5,653	-$399	-	-
Two years pre-IAH	$5,410	$5,853	-$443	-$44 ($72)	-1.0

Source: Mathematica's analysis of data from the IAH implementation contractor and Medicare claims and enrollment data from the Chronic Conditions Warehouse.

Notes: Total unweighted number of observations across all years is 259,609. The models used to estimate effects in this table are identical to those used in Exhibit B.2a with the exception of a single binary indicator for being hospitalized with COVID-19 between the time the beneficiary entered the sample and the end of the demonstration year. Standard errors are in parentheses. We computed coefficients and standard errors by using the weighted sample size, which considers the matching and eligibility weights. The table reports the regression-adjusted means of the IAH and matched comparison groups in each year. We obtained regression-adjusted means by applying the estimated regression coefficients to the covariates of IAH beneficiaries in the latest demonstration year. We calculated percentage effects using the unadjusted IAH group mean in the year before the demonstration. Because of rounding, a difference-in-differences estimate displayed as zero may be shown alongside a percentage effect that exceeds zero. Estimated effects in Year 7 are interpreted as effects of IAH during the first year of the COVID-19 pandemic.

*/**/*** The difference is statistically significant at the 0.10/0.05/0.01 level.

[a] The difference-in-differences estimate for the year before the demonstration is zero (with no standard error) in all regressions because we calculated that estimate for each year as the difference in means between IAH and comparison beneficiaries in that year minus the difference in the year before the demonstration.

PBPM = per beneficiary per month.

Exhibit B.6a. Subgroup beneficiary sample sizes for IAH and comparison beneficiaries, in Year 7

Subgroup	IAH	Comparison
ADLs		
Required assistance from another person with 2-4 ADLs	1,671 (33.0%)	7,270 (34.9%)
Required assistance from another person with 5-6 ADLs	3,387 (67.0%)	13,561 (65.1%)
Dual eligibility		
Not dually eligible	2,906 (57.5%)	12,945 (62.1%)
Dually eligible	2,152 (42.5%)	7,886 (37.9%)
Race[a]		
Non-Hispanic White	3,490 (69.0%)	14,649 (70.3%)
Non-Hispanic Black	1,193 (23.6%)	4,628 (22.2%)
Other races and ethnicities	375 (7.4%)	1,554 (7.5%)
Age		
Under 85 years	3,564 (70.5%)	14,740 (70.8%)
85 years and older	1,494 (29.5%)	6,091 (29.2%)
Chronic conditions		
Fewer than 10 chronic conditions	2,810 (55.8%)	11,706 (56.2%)
10 or more chronic conditions	2,228 (44.2%)	9,125 (43.8%)
Original reason for entitlement		
Reason for entitlement was being aged	3,024 (59.8%)	13,358 (64.1%)
Reason for entitlement was disability, ESRD, or both	2,034 (40.2%)	7,473 (35.9%)

Source: Mathematica's analysis of data from the IAH implementation contractor and Medicare claims and enrollment data from the Chronic Conditions Warehouse.

Notes: The table reports the sample size for each subgroup in Year 7 for IAH and matched comparison groups and percentage of the group in each subgroup (for example, percentage of IAH beneficiaries who required assistance from another person with 2 to 4 ADLs). As we discuss in Appendix A, all 12 subgroups had good balance between IAH and comparison beneficiaries on variables used in propensity score matching.

[a] Percents do not sum to 100% and the sum of both groups are smaller than other subgroups because beneficiaries that do not fall into either of these race groups were excluded in this analysis.

ADL = activities of daily living; ESRD = end-stage renal disease.

Exhibit B.6b. Estimated effect of IAH on Medicare spending PBPM in Years 1 to 7, by number of ADLs with which beneficiary required assistance from another person, practices that participated in Year 7

Service type	IAH	Comparison	Difference (IAH - comparison)	Difference-in-differences estimated effect	Percentage effect[a]
2-4 ADL, total spending					
Year 7	$3,859	$4,102	-$243	-$14 ($263)	-0.4
Year 6	$3,798	$4,044	-$246	-$17 ($180)	-0.5
Year 5	$3,675	$3,939	-$264	-$35 ($194)	-1.0
Year 4	$3,483	$3,549	-$66	$163 ($214)	4.8
Year 3	$3,528	$3,626	-$98	$131 ($152)	3.8
Year 2	$3,764	$3,830	-$66	$163 ($116)	4.8
Year 1	$3,553	$3,659	-$106	$123 ($79)	3.6
One year pre-IAH[b]	$3,651	$3,880	-$229	-	-
Two years pre-IAH	$3,769	$3,986	-$217	$12 ($76)	0.4
5-6 ADL, total spending					
Year 7	$5,575	$6,769	-$1,194	-$704*** ($214)	-14.0
Year 6	$5,680	$6,386	-$706	-$216 ($172)	-4.3
Year 5	$5,579	$6,253	-$674	-$184 ($194)	-3.7
Year 4	$5,374	$5,996	-$621	-$131 ($202)	-2.6
Year 3	$5,549	$5,949	-$401	$90 ($136)	1.8
Year 2	$5,649	$5,883	-$234	$256 ($163)	5.1
Year 1	$5,665	$5,995	-$330	$160 ($153)	3.2
One year pre-IAH[b]	$5,575	$6,066	-$490	-	-
Two years pre-IAH	$5,766	$6,331	-$565	-$75 ($99)	-1.5
Wald test for difference between Year 7 estimated effect by group: p<0.001					
2-4 ADL, inpatient spending					
Year 7	$1,589	$1,853	-$264	$71 ($193)	6.0
Year 6	$1,451	$1,739	-$288	$46 ($135)	3.9
Year 5	$1,405	$1,724	-$319	$15 ($128)	1.3
Year 4	$1,316	$1,450	-$134	$200 ($141)	16.9
Year 3	$1,289	$1,474	-$185	$150 ($102)	12.6
Year 2	$1,353	$1,600	-$247	$87 ($90)	7.3
Year 1	$1,289	$1,485	-$195	$139** ($69)	11.7
One year pre-IAH[b]	$1,274	$1,608	-$334	-	-
Two years pre-IAH	$1,352	$1,627	-$275	$59 ($72)	5.0

Exhibit B.6b (*continued*)

Service type	IAH	Comparison	Difference (IAH - comparison)	Difference-in-differences estimated effect	Percentage effect[a]
5-6 ADL, inpatient spending					
Year 7	$2,447	$3,282	-$835	-$431*** ($124)	-20.3
Year 6	$2,483	$3,045	-$562	-$159 ($101)	-7.5
Year 5	$2,450	$2,983	-$533	-$129 ($134)	-6.1
Year 4	$2,386	$2,830	-$444	-$40 ($125)	-1.9
Year 3	$2,466	$2,803	-$337	$67 ($106)	3.2
Year 2	$2,515	$2,778	-$263	$141 ($92)	6.6
Year 1	$2,514	$2,800	-$286	$118 ($98)	5.5
One year pre-IAH[b]	$2,427	$2,831	-$404	-	-
Two years pre-IAH	$2,491	$2,958	-$467	-$64 ($81)	-3.0
Wald test for difference between Year 7 estimated effect by group: p<0.001					

Source: Mathematica's analysis of data from the IAH implementation contractor and Medicare claims and enrollment data from the Chronic Conditions Warehouse.

Notes: Subgroup estimates are calculated using one regression that interacts a binary indicator for subgroup with all covariates in the model, once for total spending and once for inpatient spending. A Wald test was used to compare Year 7 effect estimates in each subgroup. Standard errors are in parentheses. We computed coefficients and standard errors by using the weighted sample size, which considers the matching and eligibility weights. The table reports the regression-adjusted means of the IAH and matched comparison groups in each year. We obtained regression-adjusted means by applying the estimated regression coefficients to the covariates of IAH beneficiaries in the subgroup in the latest demonstration year. Because of rounding, a difference-in-differences estimate displayed as zero may be shown alongside a percentage effect that exceeds zero. Estimated effects in Year 7 are interpreted as effects of IAH during the first year of the COVID-19 pandemic.

*/**/*** The difference is statistically significant at the 0.10/0.05/0.01 level.

[a] We used the unadjusted IAH subgroup mean in the year before the demonstration to calculate the percentage effect for each demonstration year.

[b] The difference-in-differences estimate for the period before the demonstration is zero (with no standard error) in all regressions because we calculated that estimate for each year as the difference in means between IAH and comparison beneficiaries in that year minus the difference in the year before the demonstration.

ADL = activities of daily living; PBPM = per beneficiary per month.

Exhibit B.6c. Estimated effect of IAH on Medicare spending PBPM in Years 1 to 7, by number of ADLs with which beneficiary required assistance from another person, practices that participated in Year 7, controlling for COVID-19 diagnosis and hospitalization

Service type	IAH	Comparison	Difference (IAH - comparison)	Difference-in-differences estimated effect	Percentage effect[a]
5-6 ADL, total spending, controlling for COVID-19 diagnosis in Year 7					
Year 7	$5,575	$6,715	-$1,140	-$650 (208)***	-12.9
Year 6	$6,132	$6,839	-$707	-$217 (172)	-4.3
Year 5	$6,030	$6,705	-$675	-$185 (194)	-3.7
Year 4	$5,827	$6,447	-$620	-$130 (202)	-2.6
Year 3	$5,999	$6,399	-$400	$90 (136)	1.8
Year 2	$6,100	$6,334	-$234	$256 (163)	5.1
Year 1	$6,114	$6,444	-$330	$160 (153)	3.2
One year pre-IAH[b]	$6,026	$6,516	-$490	-	-
Two years pre-IAH	$6,214	$6,780	-$566	-$76 (99)	-1.5
5-6 ADL, total spending, controlling for COVID-19 hospitalization in Year 7					
Year 7	$5,575	$6,759	-$1,184	-$694 (216)***	-13.8
Year 6	$5,999	$6,706	-$707	-$217 (172)	-4.3
Year 5	$5,897	$6,572	-$675	-$185 (194)	-3.7
Year 4	$5,693	$6,313	-$621	-$131 (202)	-2.6
Year 3	$5,866	$6,267	-$400	$90 (136)	1.8
Year 2	$5,967	$6,201	-$234	$256 (163)	5.1
Year 1	$5,982	$6,312	-$330	$160 (153)	3.2
One year pre-IAH[b]	$5,894	$6,384	-$490	. (.)	
Two years pre-IAH	$6,083	$6,649	-$566	-$75 (99)	-1.5

Source: Mathematica's analysis of data from the IAH implementation contractor and Medicare claims and enrollment data from the Chronic Conditions Warehouse.

Notes: Subgroup estimates are calculated using one regression that interacts a binary indicator for subgroup with all covariates in the model, once for total spending and once for inpatient spending. A Wald test was used to compare Year 7 effect estimates in each subgroup. Standard errors are in parentheses. We computed coefficients and standard errors by using the weighted sample size, which considers the matching and eligibility weights. The table reports the regression-adjusted means of the IAH and matched comparison groups in each year. We obtained regression-adjusted means by applying the estimated regression coefficients to the covariates of IAH beneficiaries in the subgroup in the latest demonstration year. Because of rounding, a difference-in-differences estimate displayed as zero may be shown alongside a percentage effect that exceeds zero. Estimated effects in Year 7 are interpreted as effects of IAH during the first year of the COVID-19 pandemic.

*/**/*** The difference is statistically significant at the 0.10/0.05/0.01 level.

[a] We used the unadjusted IAH subgroup mean in the year before the demonstration to calculate the percentage effect for each demonstration year.

Exhibit B.6c (*continued*)

[b] The difference-in-differences estimate for the period before the demonstration is zero (with no standard error) in all regressions because we calculated that estimate for each year as the difference in means between IAH and comparison beneficiaries in that year minus the difference in the year before the demonstration.

ADL = activities of daily living; PBPM = per beneficiary per month.

Exhibit B.6d. Estimated effect of IAH on Medicare spending PBPM in Years 1 to 7, by whether beneficiary was dually eligible for Medicare and Medicaid, practices that participated in Year 7

Service type	IAH	Comparison	Difference (IAH - comparison)	Difference-in-differences estimated effect	Percentage effect[a]
Non-dually eligible, total spending					
Year 7	$4,758	$5,462	-$704	-$321** ($144)	-7.9
Year 6	$4,962	$5,334	-$373	$11 ($158)	0.3
Year 5	$4,774	$5,208	-$434	-$51 ($161)	-1.3
Year 4	$4,616	$4,850	-$234	$149 ($193)	3.7
Year 3	$4,670	$4,924	-$254	$130 ($145)	3.2
Year 2	$4,802	$4,877	-$75	$308** ($126)	7.6
Year 1	$4,818	$4,913	-$94	$289*** ($106)	7.1
One year pre-IAH[b]	$4,639	$5,022	-$383	-	-
Two years pre-IAH	$4,898	$5,225	-$328	$56 ($101)	1.4
Dually eligible, total spending					
Year 7	$5,333	$6,389	-$1,055	-$631* ($352)	-13.8
Year 6	$5,229	$6,001	-$773	-$348 ($214)	-7.6
Year 5	$5,215	$5,843	-$628	-$203 ($223)	-4.4
Year 4	$4,973	$5,571	-$598	-$173 ($234)	-3.8
Year 3	$5,189	$5,539	-$350	$75 ($150)	1.6
Year 2	$5,368	$5,737	-$369	$56 ($175)	1.2
Year 1	$5,172	$5,587	-$515	-$91 ($151)	-2.0
One year pre-IAH[b]	$5,418	$5,843	-$424	-	-
Two years pre-IAH	$5,420	$6,031	-$611	-$186 ($116)	-4.1
Wald test for difference between Year 7 estimated effect by group: p=0.31					
Non-dually eligible, inpatient spending					
Year 7	$1,951	$2,519	-$568	-$158* ($91)	-10.7
Year 6	$2,006	$2,404	-$398	$12 ($94)	0.8
Year 5	$1,902	$2,329	-$427	-$17 ($96)	-1.1
Year 4	$1,850	$2,110	-$260	$150 ($106)	10.1
Year 3	$1,868	$2,157	-$289	$121 ($97)	8.2
Year 2	$1,912	$2,128	-$217	$193** ($76)	13.1
Year 1	$1,938	$2,120	-$182	$228*** ($58)	15.4
One year pre-IAH[b]	$1,777	$2,187	-$410	-	-
Two years pre-IAH	$1,950	$2,274	-$324	$85 ($85)	5.8

Exhibit B.6d (*continued*)

Service type	IAH	Comparison	Difference (IAH - comparison)	Difference-in-differences estimated effect	Percentage effect[a]
Dually eligible, inpatient spending					
Year 7	$2,445	$3,145	-$699	-$351 ($237)	-17.7
Year 6	$2,347	$2,888	-$541	-$193 ($154)	-9.7
Year 5	$2,391	$2,861	-$470	-$122 ($181)	-6.2
Year 4	$2,296	$2,665	-$369	-$21 ($181)	-1.1
Year 3	$2,359	$2,618	-$259	$89 ($128)	4.5
Year 2	$2,423	$2,763	-$340	$8 ($138)	0.4
Year 1	$2,319	$2,699	-$380	-$32 ($124)	-1.6
One year pre-IAH[b]	$2,416	$2,764	-$348	-	-
Two years pre-IAH	$2,333	$2,832	-$500	-$152 ($103)	-7.7
Wald test for difference between Year 7 estimated effect by group: p=0.33					

Source: Mathematica's analysis of data from the IAH implementation contractor and Medicare claims and enrollment data from the Chronic Conditions Warehouse.

Notes: Subgroup estimates are calculated using a one regression that interacts a binary indicator for subgroup with all covariates in the model, once for total spending and once for inpatient spending. A Wald test was used to compare Year 7 effect estimates in each subgroup. Standard errors are in parentheses. We computed coefficients and standard errors by using the weighted sample size, which considers the matching and eligibility weights. The table reports the regression-adjusted means of the IAH and matched comparison groups in each year. We obtained regression-adjusted means by applying the estimated regression coefficients to the covariates of IAH beneficiaries in the subgroup in the latest demonstration year. Because of rounding, a difference-in-differences estimate displayed as zero may be shown alongside a percentage effect that exceeds zero. Estimated effects in Year 7 are interpreted as effects of IAH during the first year of the COVID-19 pandemic.

*/**/*** The difference is statistically significant at the 0.10/0.05/0.01 level.

[a] We used the unadjusted IAH subgroup mean in the year before the demonstration to calculate the percentage effect for each demonstration year.

[b] The difference-in-differences estimate for the period before the demonstration is zero (with no standard error) in all regressions because we calculated that estimate for each year as the difference in means between IAH and comparison beneficiaries in that year minus the difference in the year before the demonstration.

PBPM = per beneficiary per month.

Exhibit B.6e. Estimated effect of IAH on Medicare spending PBPM in Years 1 to 7, by race, practices that participated in Year 7

Service type	IAH	Comparison	Difference (IAH - comparison)	Difference-in-differences estimated effect	Percentage effect[a]
Non-Hispanic White, total spending					
Year 7	$4,679	$5,415	-$736	-$438*** ($128)	-10.5
Year 6	$4,832	$5,243	-$411	-$113 ($126)	-2.7
Year 5	$4,776	$5,173	-$398	-$99 ($152)	-2.4
Year 4	$4,452	$4,784	-$332	-$33 ($161)	-0.8
Year 3	$4,640	$4,888	-$248	$50 ($105)	1.2
Year 2	$4,724	$4,863	-$139	$159 ($119)	3.8
Year 1	$4,744	$4,926	-$182	$116 ($87)	2.8
One year pre-IAH[b]	$4,740	$5,039	-$298	-	-
Two years pre-IAH	$4,860	$5,242	-$382	-$84 ($89)	-2.0
Non-Hispanic Black, total spending					
Year 7	$5,750	$7,108	-$1,358	-$704 ($447)	-15.2
Year 6	$5,646	$6,616	-$970	-$316 ($311)	-6.8
Year 5	$5,429	$6,330	-$901	-$247 ($296)	-5.3
Year 4	$5,634	$6,138	-$504	$151 ($317)	3.3
Year 3	$5,645	$6,040	-$395	$259 ($229)	5.6
Year 2	$5,923	$6,231	-$308	$346 ($262)	7.5
Year 1	$5,627	$6,141	-$514	$141 ($211)	3.0
One year pre-IAH[b]	$5,694	$6,348	-$654	-	-
Two years pre-IAH	$5,963	$6,485	-$522	$132 ($200)	2.9
Wald test for difference between Year 7 estimated effect versus the estimate for non-Hispanic White: p=0.48					
Other races and ethnicities					
Year 7	$5,602	$5,976	-$374	$66 ($527)	1.7
Year 6	$5,534	$5,926	-$392	$48 ($667)	1.3
Year 5	$5,390	$5,694	-$304	$136 ($473)	3.6
Year 4	$4,849	$5,460	-$611	-$171 ($643)	-4.5
Year 3	$4,858	$5,302	-$444	-$4 ($602)	-0.1
Year 2	$5,371	$5,706	-$334	$105 ($585)	2.8
Year 1	$5,475	$5,382	$92	$532 ($517)	14.0
One year pre-IAH[b]	$4,938	$5,377	-$440	-	-
Two years pre-IAH	$4,824	$5,846	-$1022	-$583 ($368)	-15.3
Wald test for difference between Year 7 estimated effect versus the estimate for non-Hispanic White: p=0.37					
Non-Hispanic White, inpatient spending					
Year 7	$1,928	$2,473	-$546	-$241*** ($85)	-14.8

Exhibit B.6e (*continued*)

Service type	IAH	Comparison	Difference (IAH - comparison)	Difference-in-differences estimated effect	Percentage effect[a]
Year 6	$1,990	$2,367	-$376	-$72 ($85)	-4.4
Year 5	$1,954	$2,347	-$393	-$89 ($99)	-5.5
Year 4	$1,809	$2,101	-$292	$13 ($95)	0.8
Year 3	$1,892	$2,156	-$264	$40 ($69)	2.5
Year 2	$1,893	$2,141	-$248	$57 ($74)	3.5
Year 1	$1,932	$2,163	-$231	$74 ($64)	4.5
One year pre-IAH[b]	$1,918	$2,223	-$305	-	-
Two years pre-IAH	$1,963	$2,307	-$344	-$40 ($73)	-2.5
Non-Hispanic Black, inpatient spending					
Year 7	$2,717	$3,687	-$970	-$371 ($309)	-19.7
Year 6	$2,518	$3,260	-$742	-$144 ($207)	-7.6
Year 5	$2,454	$3,125	-$671	-$73 ($204)	-3.9
Year 4	$2,642	$2,999	-$358	$241 ($206)	12.8
Year 3	$2,605	$2,926	-$321	$278 ($183)	14.7
Year 2	$2,781	$3,073	-$292	$306* ($170)	16.2
Year 1	$2,600	$2,952	-$352	$246* ($144)	13.0
One year pre-IAH[b]	$2,469	$3,068	-$598	-	-
Two years pre-IAH	$2,625	$3,097	-$473	$126 ($165)	6.7
Wald test for difference between Year 7 estimated effect versus the estimate for non-Hispanic White: p=0.62					
Other races and ethnicities					
Year 7	$2,541	$2,850	-$309	17 (354)	1.2
Year 6	$2,481	$2,828	-$347	-20 (450)	-1.3
Year 5	$2,574	$2,773	-$199	128 (408)	8.5
Year 4	$2,181	$2,573	-$392	-65 (516)	-4.3
Year 3	$2,153	$2,436	-$283	44 (443)	2.9
Year 2	$2,334	$2,757	-$424	-97 (476)	-6.4
Year 1	$2,361	$2,485	-$124	203 (414)	13.5
One year pre-IAH[b]	$2,106	$2,433	-$327	-	-
Two years pre-IAH	$1,920	$2,746	-$826	-499 (294)*	-33.2
Wald test for difference between Year 7 estimated effect versus the estimate for non-Hispanic White: p=0.52					

Source: Mathematica's analysis of data from the IAH implementation contractor and Medicare claims and enrollment data from the Chronic Conditions Warehouse.

Notes: Subgroup estimates are calculated using a one regression that interacts a binary indicator for subgroup with all covariates in the model, once for total spending and once for inpatient spending. A Wald test was used to compare Year 7 effect estimates in each subgroup. Standard errors are in parentheses. We computed coefficients and standard errors by using the weighted sample size, which considers the matching and eligibility weights. The table reports the regression-adjusted means of the IAH and matched

Exhibit B.6e (*continued*)

comparison groups in each year. We obtained regression-adjusted means by applying the estimated regression coefficients to the covariates of IAH beneficiaries in the subgroup in the latest demonstration year. Because of rounding, a difference-in-differences estimate displayed as zero may be shown alongside a percentage effect that exceeds zero. Estimated effects in Year 7 are interpreted as effects of IAH during the first year of the COVID-19 pandemic.

*/**/*** The difference is statistically significant at the 0.10/0.05/0.01 level.

[a] We used the unadjusted IAH subgroup mean in the year before the demonstration to calculate the percentage effect for each demonstration year.

[b] The difference-in-differences estimate for the period before the demonstration is zero (with no standard error) in all regressions because we calculated that estimate for each year as the difference in means between IAH and comparison beneficiaries in that year minus the difference in the year before the demonstration.

PBPM = per beneficiary per month.

Exhibit B.6f. Estimated effect of IAH on Medicare spending PBPM in Years 1 to 7, by age, practices that participated in Year 7

Service type	IAH	Comparison	Difference (IAH - comparison)	Difference-in-differences estimated effect	Percentage effect[a]
Under 85, total spending					
Year 7	$5,530	$6,441	-$911	-$484* ($247)	-10.4
Year 6	$5,467	$6,081	-$614	-$186 ($169)	-4.0
Year 5	$5,387	$5,945	-$558	-$131 ($182)	-2.8
Year 4	$5,183	$5,555	-$372	$55 ($221)	1.2
Year 3	$5,289	$5,621	-$333	$95 ($149)	2.0
Year 2	$5,521	$5,687	-$167	$261* ($136)	5.6
Year 1	$5,372	$5,675	-$304	$124 ($118)	2.7
One year pre-IAH[b]	$5,409	$5,837	-$427	-	
Two years pre-IAH	$5,561	$6,047	-$486	-$59 ($108)	-1.3
85 or over, total spending					
Year 7	$3,693	$4,397	-$704	-$370* ($200)	-10.5
Year 6	$4,082	$4,474	-$393	-$59 ($189)	-1.7
Year 5	$3,898	$4,311	-$413	-$79 ($200)	-2.2
Year 4	$3,693	$4,132	-$439	-$105 ($193)	-3.0
Year 3	$3,860	$4,058	-$198	$136 ($135)	3.8
Year 2	$3,820	$4,085	-$265	$69 ($151)	2.0
Year 1	$3,916	$4,107	-$191	$143 ($137)	4.0
One year pre-IAH[b]	$3,841	$4,175	-$334	-	-
Two years pre-IAH	$4,004	$4,350	-$346	-$12 ($100)	-0.3
Wald test for difference between Year 7 estimated effect by group: p=0.67					
Under 85, inpatient spending					
Year 7	$2,482	$3,159	-$677	-$253 ($180)	-13.1
Year 6	$2,389	$2,909	-$519	-$95 ($116)	-4.9
Year 5	$2,386	$2,864	-$478	-$54 ($127)	-2.8
Year 4	$2,297	$2,614	-$317	$107 ($142)	5.6
Year 3	$2,318	$2,637	-$319	$105 ($113)	5.5
Year 2	$2,416	$2,696	-$279	$145 ($93)	7.5
Year 1	$2,364	$2,643	-$279	$146 ($91)	7.6
One year pre-IAH[b]	$2,308	$2,733	-$424	-	-
Two years pre-IAH	$2,378	$2,821	-$443	-$19 ($89)	-1.3

Exhibit B.6f (*continued*)

Service type	IAH	Comparison	Difference (IAH - comparison)	Difference-in-differences estimated effect	Percentage effect[a]
85 or over, inpatient spending					
Year 7	$1,363	$1,857	-$494	-$200** ($100)	-16.5
Year 6	$1,543	$1,874	-$331	-$37 ($108)	-3.0
Year 5	$1,421	$1,784	-$363	-$69 ($147)	-5.7
Year 4	$1,363	$1,663	-$300	-$6 ($120)	-0.5
Year 3	$1,446	$1,628	-$182	$112 ($77)	9.3
Year 2	$1,387	$1,632	-$245	$49 ($94)	4.1
Year 1	$1,417	$1,641	-$224	$70 ($82)	5.8
One year pre-IAH[b]	$1,375	$1,669	-$294	-	-
Two years pre-IAH	$1,444	$1,735	-$291	$3 ($67)	0.2
Wald test for difference between Year 7 estimated effect by group: p=0.79					

Source: Mathematica's analysis of data from the IAH implementation contractor and Medicare claims and enrollment data from the Chronic Conditions Warehouse.

Notes: Subgroup estimates are calculated using a one regression that interacts a binary indicator for subgroup with all covariates in the model, once for total spending and once for inpatient spending. A Wald test was used to compare Year 7 effect estimates in each subgroup. Standard errors are in parentheses. We computed coefficients and standard errors by using the weighted sample size, which considers the matching and eligibility weights. The table reports the regression-adjusted means of the IAH and matched comparison groups in each year. We obtained regression-adjusted means by applying the estimated regression coefficients to the covariates of IAH beneficiaries in the subgroup in the latest demonstration year. Because of rounding, a difference-in-differences estimate displayed as zero may be shown alongside a percentage effect that exceeds zero. Estimated effects in Year 7 are interpreted as effects of IAH during the first year of the COVID-19 pandemic.

*/**/*** The difference is statistically significant at the 0.10/0.05/0.01 level.

[a] We used the unadjusted IAH subgroup mean in the year before the demonstration to calculate the percentage effect for each demonstration year.

[b] The difference-in-differences estimate for the period before the demonstration is zero (with no standard error) in all regressions because we calculated that estimate for each year as the difference in means between IAH and comparison beneficiaries in that year minus the difference in the year before the demonstration.

PBPM = per beneficiary per month.

Exhibit B.6g. Estimated effect of IAH on Medicare spending PBPM in Years 1 to 7, by number of chronic conditions, practices that participated in Year 7

Service type	IAH	Comparison	Difference (IAH - comparison)	Difference-in-differences estimated effect	Percentage effect[a]
Fewer than 10 chronic conditions, total spending					
Year 7	$3,576	$4,426	-$850	-$532*** ($160)	-16.7
Year 6	$3,775	$4,225	-$450	-$132 ($134)	-4.1
Year 5	$3,761	$4,217	-$456	-$137 ($148)	-4.3
Year 4	$3,783	$4,061	-$278	$40 ($218)	1.3
Year 3	$3,727	$3,928	-$201	$117 ($164)	3.7
Year 2	$3,708	$3,867	-$159	$160 ($109)	5.0
Year 1	$3,598	$3,837	-$238	$80 ($101)	2.5
One year pre-IAH[b]	$3,659	$3,977	-$318	-	-
Two years pre-IAH	$3,863	$4,144	-$281	$37 ($74)	1.2
Number of chronic conditions 10 or more, total spending					
Year 7	$6,687	$7,568	-$881	-$391 ($301)	-6.8
Year 6	$6,623	$7,270	-$647	-$157 ($209)	-2.7
Year 5	$6,373	$6,997	-$624	-$134 ($241)	-2.3
Year 4	$6,022	$6,489	-$467	$23 ($204)	0.4
Year 3	$6,269	$6,642	-$373	$117 ($148)	2.0
Year 2	$6,617	$6,868	-$251	$239 ($200)	4.1
Year 1	$6,628	$6,879	-$251	$238 ($160)	4.1
One year pre-IAH[b]	$6,540	$7,030	-$490	-	-
Two years pre-IAH	$6,652	$7,284	-$632	-$142 ($138)	-2.5
Wald test for difference between Year 7 estimated effect by group: p=0.53					
Fewer than 10 chronic conditions, inpatient spending					
Year 7	$1,313	$1,919	-$606	-$284*** ($101)	-26.9
Year 6	$1,422	$1,781	-$359	-$36 ($89)	-3.4
Year 5	$1,446	$1,820	-$374	-$51 ($114)	-4.9
Year 4	$1,478	$1,724	-$246	$76 ($125)	7.2
Year 3	$1,444	$1,626	-$181	$141 ($120)	13.4
Year 2	$1,356	$1,590	-$234	$89 ($81)	8.4
Year 1	$1,283	$1,517	-$235	$88 ($59)	8.3
One year pre-IAH[b]	$1,290	$1,613	-$323	-	-
Two years pre-IAH	$1,407	$1,696	-$288	$34 ($64)	3.2

Exhibit B.6g (*continued*)

Service type	IAH	Comparison	Difference (IAH - comparison)	Difference-in-differences estimated effect	Percentage effect[a]
Number of chronic conditions 10 or more, inpatient spending					
Year 7	$3,163	$3,829	-$666	-$217 ($202)	-8.5
Year 6	$3,018	$3,595	-$576	-$127 ($143)	-5.0
Year 5	$2,882	$3,444	-$562	-$113 ($151)	-4.4
Year 4	$2,754	$3,107	-$353	$97 ($139)	3.8
Year 3	$2,816	$3,195	-$378	$71 ($107)	2.8
Year 2	$3,037	$3,359	-$322	$127 ($131)	5.0
Year 1	$3,097	$3,360	-$263	$186 ($119)	7.3
One year pre-IAH[b]	$2,963	$3,412	-$449	-	-
Two years pre-IAH	$2,980	$3,500	-$520	-$71 ($124)	-2.8
Wald test for difference between Year 7 estimated effect by group: p=0.65					

Source: Mathematica's analysis of data from the IAH implementation contractor and Medicare claims and enrollment data from the Chronic Conditions Warehouse.

Notes: Subgroup estimates are calculated using a one regression that interacts a binary indicator for subgroup with all covariates in the model, once for total spending and once for inpatient spending. A Wald test was used to compare Year 7 effect estimates in each subgroup. Standard errors are in parentheses. We computed coefficients and standard errors by using the weighted sample size, which considers the matching and eligibility weights. The table reports the regression-adjusted means of the IAH and matched comparison groups in each year. We obtained regression-adjusted means by applying the estimated regression coefficients to the covariates of IAH beneficiaries in the subgroup in the latest demonstration year. Because of rounding, a difference-in-differences estimate displayed as zero may be shown alongside a percentage effect that exceeds zero. Estimated effects in Year 7 are interpreted as effects of IAH during the first year of the COVID-19 pandemic.

*/**/*** The difference is statistically significant at the 0.10/0.05/0.01 level.

[a] We used the unadjusted IAH subgroup mean in the year before the demonstration to calculate the percentage effect for each demonstration year.

[b] The difference-in-differences estimate for the period before the demonstration is zero (with no standard error) in all regressions because we calculated that estimate for each year as the difference in means between IAH and comparison beneficiaries in that year minus the difference in the year before the demonstration.

PBPM = per beneficiary per month.

Exhibit B.6h. Estimated effect of IAH on Medicare spending PBPM in Years 1 to 7, by original reason for Medicare entitlement, practices that participated in Year 7

Service type	IAH	Comparison	Difference (IAH - comparison)	Difference-in-differences estimated effect	Percentage effect[a]
Original reason for Medicare entitlement was being aged, total spending					
Year 7	$4,444	$5,229	-$785	-$503*** ($159)	-12.5
Year 6	$4,685	$5,105	-$420	-$137 ($160)	-3.4
Year 5	$4,528	$4,963	-$435	-$153 ($196)	-3.8
Year 4	$4,302	$4,647	-$345	-$62 ($207)	-1.6
Year 3	$4,415	$4,641	-$227	$56 ($130)	1.4
Year 2	$4,496	$4,655	-$159	$123 ($150)	3.1
Year 1	$4,480	$4,716	-$235	$47 ($99)	1.2
One year pre-IAH[b]	$4,494	$4,777	-$282	-	-
Two years pre-IAH	$4,633	$4,995	-$363	-$80 ($78)	-2.0
Original reason for Medicare entitlement was having a disability, ESRD, or both, total spending					
Year 7	$5,819	$6,779	-$960	-$346 ($323)	-7.2
Year 6	$5,632	$6,363	-$730	-$115 ($203)	-2.4
Year 5	$5,582	$6,225	-$644	-$29 ($202)	-0.6
Year 4	$5,427	$5,898	-$470	$144 ($228)	3.0
Year 3	$5,552	$5,961	-$410	$205 ($218)	4.3
Year 2	$5,807	$6,103	-$296	$318* ($172)	6.7
Year 1	$5,653	$5,978	-$325	$290 ($176)	6.1
One year pre-IAH[b]	$5,625	$6,240	-$615	-	-
Two years pre-IAH	$5,789	$6,392	-$603	$12 ($151)	0.3
Wald test for difference between Year 7 estimated effect by group: p=0.55					
Original reason for Medicare entitlement was being aged, inpatient spending					
Year 7	$1,823	$2,374	-$551	-$262*** ($95)	-17.5
Year 6	$1,903	$2,265	-$361	-$72 ($91)	-4.8
Year 5	$1,801	$2,206	-$405	-$116 ($122)	-7.7
Year 4	$1,718	$1,986	-$267	$22 ($122)	1.5
Year 3	$1,761	$1,987	-$226	$63 ($85)	4.2
Year 2	$1,772	$1,998	-$225	$64 ($92)	4.3
Year 1	$1,764	$2,008	-$244	$45 ($68)	3.0
One year pre-IAH[b]	$1,734	$2,024	-$289	-	-
Two years pre-IAH	$1,776	$2,124	-$348	-$58 ($61)	-3.9

Exhibit B.6h (*continued*)

Service type	IAH	Comparison	Difference (IAH - comparison)	Difference-in-differences estimated effect	Percentage effect[a]
Original reason for Medicare entitlement was having a disability, ESRD, or both, inpatient spending					
Year 7	$2,655	$3,398	-$744	-$186 ($241)	-9.1
Year 6	$2,510	$3,118	-$609	-$51 ($152)	-2.5
Year 5	$2,558	$3,067	-$508	$49 ($172)	2.4
Year 4	$2,495	$2,880	-$386	$172 ($170)	8.4
Year 3	$2,514	$2,885	-$371	$187 ($177)	9.1
Year 2	$2,623	$2,997	-$374	$184 ($134)	9.0
Year 1	$2,575	$2,874	-$299	$259* ($133)	12.6
Wald test for difference between Year 7 estimated effect by group: p=0.72					

Source: Mathematica's analysis of data from the IAH implementation contractor and Medicare claims and enrollment data from the Chronic Conditions Warehouse.

Notes: Subgroup estimates are calculated using a one regression that interacts a binary indicator for subgroup with all covariates in the model, once for total spending and once for inpatient spending. A Wald test was used to compare Year 7 effect estimates in each subgroup. Standard errors are in parentheses. We computed coefficients and standard errors by using the weighted sample size, which considers the matching and eligibility weights. The table reports the regression-adjusted means of the IAH and matched comparison groups in each year. We obtained regression-adjusted means by applying the estimated regression coefficients to the covariates of IAH beneficiaries in the subgroup in the latest demonstration year. Because of rounding, a difference-in-differences estimate displayed as zero may be shown alongside a percentage effect that exceeds zero. Estimated effects in Year 7 are interpreted as effects of IAH during the first year of the COVID-19 pandemic.

*/**/*** The difference is statistically significant at the 0.10/0.05/0.01 level.

[a] We used the unadjusted IAH subgroup mean in the year before the demonstration to calculate the percentage effect for each demonstration year.

[b] The difference-in-differences estimate for the period before the demonstration is zero (with no standard error) in all regressions because we calculated that estimate for each year as the difference in means between IAH and comparison beneficiaries in that year minus the difference in the year before the demonstration.

PBPM = per beneficiary per month.

Exhibit B.7. Estimated effects of IAH on total Medicare spending PBPM excluding one site at a time, average annual effects and effects in Year 7

Excluding site	Estimated effect in Year 7	Percentage effect in Year 7[a]	Estimated average annual effect across all 7 years
A	-$646*** ($178)	-14.7	-$288** ($136)
B	-$480** ($218)	-11.1	-$202 ($141)
C	-$476** ($222)	-11.1	-$176 ($143)
D	-$475** ($210)	-11.0	-$204 ($137)
E	-$469** ($208)	-10.9	-$205 ($136)
F	-$466** ($213)	-10.9	-$207 ($138)
G	-$463** ($211)	-10.8	-$207 ($138)
H	-$458** ($207)	-10.7	-$203 ($136)
I	-$455** ($212)	-10.6	-$201 ($139)
J	-$431** ($219)	-10.2	-$272* ($151)
K	-$379** ($181)	-9.0	-$166 ($143)
L	-$326 ($226)	-7.7	-$199* ($118)
M	n.a.	n.a.	-$221 ($145)
N	n.a.	n.a.	-$199 ($138)
O	n.a.	n.a.	-$196 ($138)
P	n.a.	n.a.	-$44 ($121)
Full sample	**-$459** ($206)**	**-10.7**	**-$200 ($135)**

Source: Mathematica's analysis of data from the IAH implementation contractor and Medicare claims and enrollment data from the Chronic Conditions Warehouse.

Notes: For each outcome in the table, we estimated 12 to 16 regressions (treating each of the three organizations in the Richmond-based consortium separately), with each regression excluding the IAH beneficiaries from one practice and their matched comparisons in all years. Standard errors are in parentheses. We computed coefficients and standard errors by using the weighted sample size, which considers the matching and eligibility weights. Estimated effects in Year 7 are interpreted as effects of IAH during the first year of the COVID-19 pandemic. We estimated average annual effects and their standard errors using seemingly unrelated regression (see Chapter 6 of Appendix A).

*/**/*** The difference is statistically significant at the 0.10/0.05/0.01 level.

[a] We used the unadjusted IAH group mean in the year before the demonstration excluding the respective site to calculate the percentage effect for each demonstration year.

PBPM = per beneficiary per month; n.a. = not applicable.

Exhibit B.8a. Estimated effect of IAH on total Medicare spending PBPM for IAH beneficiaries at practices that participated in Year 7, using beneficiary and practice weighting

	Beneficiary weighting		Practice weighting	
	Estimated effect	Percentage effect[a]	Estimated effect	Percentage effect[a]
Year 7	-$459** ($206)	-10.7	-$422*** (159)	-9.9
Year 6	-$148 ($150)	-3.4	$75 ($179)	1.8
Year 5	-$123 ($179)	-2.9	-$50 ($152)	-1.2
Year 4	$11 ($202)	0.2	-$25 ($133)	-0.6
Year 3	$112 ($129)	2.6	$57 ($112)	1.3
Year 2	$203 ($127)	4.7	$209* ($125)	4.9
Year 1	$136 ($103)	3.2	$23 ($148)	0.5
One year pre-IAH[b]	-	-	-	-
Two years pre-IAH	-$44 ($72)	-1.0	-$12 ($123)	-0.3

Source: Mathematica's analysis of data from the IAH implementation contractor and Medicare claims and enrollment data from the Chronic Conditions Warehouse.

Notes: Total unweighted number of observations across all years is 259,609. Standard errors are in parentheses. We computed coefficients and standard errors by using the weighted sample size, which considers the matching and eligibility weights. The table reports the regression-adjusted means of the IAH and matched comparison groups in each year. We obtained regression-adjusted means by applying the estimated regression coefficients to the covariates of IAH beneficiaries in the latest demonstration year. Because of rounding, a difference-in-differences estimate displayed as zero may be shown alongside a percentage effect that exceeds zero. Beneficiary weighting treats each beneficiary equally regardless of which practice they were affiliated with; under this approach, larger IAH practices typically influence the estimated effect more than smaller practices. Practice weighting treates each IAH practice equally regardless of how many IAH-eligible patients they treated. For more information about the difference between beneficiary and practice weighting, see Appendix A, Section 6. Estimated effects in Year 7 are interpreted as effects of IAH during the first year of the COVID-19 pandemic.

*/**/*** The difference is statistically significant at the 0.10/0.05/0.01 level.

[a] We used the unadjusted IAH group mean in the year before the demonstration to calculate the percentage effect for each demonstration year. Exhibit B.1 reports the baseline unadjusted IAH group mean for all outcomes.

[b] The difference-in-differences estimate for the period before the demonstration is zero (with no standard error) in all regressions because we calculated that estimate for each year as the difference in means between IAH and comparison beneficiaries in that year minus the difference in the year before the demonstration.

PBPM = per beneficiary per month.

Exhibit B.8b. Estimated effect of the IAH payment incentive on total Medicare spending PBPM for IAH beneficiaries at practices that participated in Year 6, using beneficiary and practice weighting

	Beneficiary weighting		Practice weighting	
	Estimated effect	Percentage effect[a]	Estimated effect	Percentage effect[a]
Year 6	-$41 ($126)	-1.0	$76 ($166)	1.8
Year 5	-$113 ($157)	-2.7	-$59 ($143)	-1.4
Year 4	-$19 ($179)	-0.4	-$59 ($136)	-1.4
Year 3	$52 ($117)	1.3	$21 ($109)	0.5
Year 2	$150 ($122)	3.6	$176 ($116)	4.1
Year 1	$101 ($92)	2.4	-$15 ($138)	-0.4
One year pre-IAH[b]	-	-	-	-
Two years pre-IAH	-$27 ($64)	-0.6	$22 ($113)	0.5

Source: Mathematica's analysis of data from the IAH implementation contractor and Medicare claims and enrollment data from the Chronic Conditions Warehouse.

Notes: Total unweighted number of observations across all years is 290,514. Standard errors are in parentheses. We computed coefficients and standard errors by using the weighted sample size, which considers the matching and eligibility weights. The table reports the regression-adjusted means of the IAH and matched comparison groups in each year. We obtained regression-adjusted means by applying the estimated regression coefficients to the covariates of IAH beneficiaries in the latest demonstration year. Because of rounding, a difference-in-differences estimate displayed as zero may be shown alongside a percentage effect that exceeds zero. For information about the difference between beneficiary and practice weighting, see Appendix A, Section 6.

*/**/*** The difference is statistically significant at the 0.10/0.05/0.01 level.

[a] We used the unadjusted IAH group mean in the year before the demonstration to calculate the percentage effect for each demonstration year. Exhibit B.1 reports the baseline unadjusted IAH group mean for all outcomes.

[b] The difference-in-differences estimate for the period before the demonstration is zero (with no standard error) in all regressions because we calculated that estimate for each year as the difference in means between IAH and comparison beneficiaries in that year minus the difference in the year before the demonstration.

PBPM = per beneficiary per month.

Exhibit B.9. Estimated effect of IAH on total Medicare spending PBPM in Year 7, with and without MIPS adjustments, practices that participated in Year 7

	IAH	Comparison	Difference (IAH - comparison)	Difference-in-differences estimated effect	Percentage effect[a]
Year 7 (including MIPS adjustments)[b]	$5,002	$5,862	-$860	-$459** ($206)	-10.7
Year 7 (excluding MIPS adjustments)	$4,999	$5,858	-$860	-$459** ($206)	-10.7

Source: Mathematica's analysis of data from the IAH implementation contractor and Medicare claims and enrollment data from the Chronic Conditions Warehouse.

Note: Standard errors are in parentheses. We computed coefficients and standard errors by using the weighted sample size, which considers the matching and eligibility weights. The table reports the regression-adjusted means of the IAH and matched comparison groups in each year. We obtained regression-adjusted means by applying the estimated regression coefficients to the covariates of IAH beneficiaries in the latest demonstration year. Because of rounding, a difference-in-differences estimate displayed as zero may be shown alongside a percentage effect that exceeds zero. Under MIPS, CMS makes payment adjustments to individual providers nationally on the basis of quality and efficiency metrics. Estimated effects in Year 7 are interpreted as effects of IAH during the first year of the COVID-19 pandemic.

[a] We used the unadjusted IAH group mean in the year before the demonstration to calculate the percentage effect for each demonstration year. Exhibit B.1 reports the baseline unadjusted IAH group mean for all outcomes.

[b] This result is the same as we reported in Exhibit B.2a.

CMS = Centers for Medicare & Medicaid Services; MIPS = Merit-based Incentive Payment System; PBPM = per beneficiary per month.

Exhibit B.10a. Estimated effect of IAH on total Medicare spending PBPM in Year 7, with and without adjusting for ACO participation, practices that participated in Year 7

	IAH	Comparison	Difference (IAH - comparison)	Difference-in-differences estimated effect	Percentage effect[a]
Year 7, without controlling for ACO participation[b]	$5,002	$5,862	-$860	-$459** ($206)	-10.7
Year 7, controlling for ACO participation	$5,002	$5,756	-$754	-$347* ($199)	-8.1

Source: Mathematica's analysis of data from the IAH implementation contractor and Medicare claims and enrollment data from the Chronic Conditions Warehouse.

Note: Standard errors are in parentheses. We computed coefficients and standard errors by using the weighted sample size, which considers the matching and eligibility weights. The table reports the regression-adjusted means of the IAH and matched comparison groups in each year. We obtained regression-adjusted means by applying the estimated regression coefficients to the covariates of IAH beneficiaries in the latest demonstration year. Because of rounding, a difference-in-differences estimate displayed as zero may be shown alongside a percentage effect that exceeds zero. ACOs are groups of providers that coordinate the care of an assigned Medicare population and are held financially accountable for the quality, cost, and experience of care they provide. For more information on the ACO analysis, see Chapter 7 of Appendix A. Estimated effects in Year 7 are interpreted as effects of IAH during the first year of the COVID-19 pandemic.

[a] We used the unadjusted IAH group mean in the year before the demonstration to calculate the percentage effect for each demonstration year. Exhibit B.1 reports the baseline unadjusted IAH group mean for all outcomes.

[b] This result is the same as we reported in Exhibit B.2a.

ACO = accountable care organization; PBPM = per beneficiary per month.

Exhibit B.10b. Estimated effect of the IAH payment incentive on total Medicare spending PBPM in Year 6, with and without adjusting for ACO participation, practices that participated in Year 6

	IAH	Comparison	Difference (IAH - comparison)	Difference-in-differences estimated effect	Percentage effect[a]
Year 6, without controlling for ACO participation[b]	$4,725	$5,137	-$412	-$41 ($126)	-1.0
Year 6, controlling for ACO participation	$4,725	$5,126	-$401	-$29 ($124)	-0.7

Source: Mathematica's analysis of data from the IAH implementation contractor and Medicare claims and enrollment data from the Chronic Conditions Warehouse

Note: ACOs are groups of providers that coordinate the care of an assigned Medicare population and are held financially accountable for the quality, cost, and experience of care they provide. In 2019, there were 487 ACO organizations with 10.4 million assigned Medicare beneficiaries.

[a] We used the unadjusted IAH group mean in the year before the demonstration to calculate the percentage effect for each demonstration year. Exhibit B.1 reports the baseline unadjusted IAH group mean for all outcomes.

[b] This result is the same as we reported in Exhibit B.2b.

ACO = accountable care organization; PBPM = per beneficiary per month.

Exhibit B.11a. Estimated effect of IAH on hospital care use in Years 1 to 7, practices that participated in Year 7

Outcome	IAH	Comparison	Difference (IAH - comparison)	Difference-in-differences estimated effect	Percentage effect[a]
Number of hospital admissions per 1,000 beneficiaries per year[b]					
Year 7	1,639	2,025	-386	-33 (81)	-1.9
Year 6	1,886	2,188	-301	52 (73)	2.9
Year 5	1,903	2,264	-361	-8 (59)	-0.5
Year 4	1,763	2,056	-293	60 (78)	3.4
Year 3	1,855	2,145	-290	63 (76)	3.6
Year 2	1,891	2,159	-268	85 (55)	4.8
Year 1	1,906	2,171	-265	88** (42)	5.0
One year pre-IAH[c]	1,942	2,295	-353	-	-
Two years pre-IAH	2,028	2,376	-348	5 (34)	0.3
Number of hospital admissions preceded by an ED visit per 1,000 beneficiaries per year					
Year 7	1,372	1,634	-262	-107 (77)	-7.4
Year 6	1,522	1,692	-169	-15 (66)	-1.0
Year 5	1,590	1,793	-204	-49 (61)	-3.4
Year 4	1,443	1,612	-169	-14 (63)	-1.0
Year 3	1,531	1,676	-145*	9 (57)	0.6
Year 2	1,552	1,664	-112	43 (50)	3.0
Year 1	1,584	1,705	-120	34 (39)	2.4
One year pre-IAH[c]	1,633	1,787	-155	-	-
Two years pre-IAH	1,727	1,860	-134	21 (31)	1.5
Number of outpatient ED visits per 1,000 beneficiaries per year[d]					
Year 7	1,332	1,450	-118	-54 (66)	-3.8
Year 6	1,645	1,797	-152	-89 (95)	-6.3
Year 5	1,573	1,702	-129	-65 (89)	-4.6
Year 4	1,537	1,622	-86	-22 (67)	-1.5
Year 3	1,685	1,723	-39	25 (87)	1.8
Year 2	1,585	1,552	33	97 (66)	6.8
Year 1	1,446	1,475	-29	34 (62)	2.4
One year pre-IAH[c]	1,458	1,522	-64	-	-
Two years pre-IAH	1,413	1,494	-81	-18 (45)	-1.2

Source: Mathematica's analysis of data from the IAH implementation contractor and Medicare claims and enrollment data from the Chronic Conditions Warehouse.

Notes: Total unweighted number of observations across all years is 259,609. Standard errors are in parentheses. We computed coefficients and standard errors by using the weighted sample size, which considers the

Exhibit B.11a (*continued*)

matching and eligibility weights. The table reports the regression-adjusted means of the IAH and matched comparison groups in each year. We obtained regression-adjusted means by applying the estimated regression coefficients to the covariates of IAH beneficiaries in the latest demonstration year. Because of rounding, a difference-in-differences estimate displayed as zero may be shown alongside a percentage effect that exceeds zero. Estimated effects in Year 7 are interpreted as effects of IAH during the first year of the COVID-19 pandemic.

*/**/*** The difference is statistically significant at the 0.10/0.05/0.01 level.

[a] We used the unadjusted IAH group mean in the year before the demonstration to calculate the percentage effect for each demonstration year. Exhibit B.1 reports the baseline unadjusted IAH group mean for all outcomes.

[b] The number of hospital admissions includes observation stays.

[c] The difference-in-differences estimate for the year before the demonstration is zero (with no standard error) in all regressions because we calculated that estimate for each year as the difference in means between IAH and comparison beneficiaries in that year minus the difference in the year before the demonstration.

[d] The number of outpatient ED visits reflects all those not resulting in a hospital admission, including those resulting in an observation stay.

ED = emergency department.

Exhibit B.11b. Estimated effect of the IAH payment incentive on hospital care use in Years 1 to 6, practices that participated in Year 6

Outcome	IAH	Comparison	Difference (IAH - comparison)	Difference-in-differences estimated effect	Percentage effect[a]
Number of hospital admissions per 1,000 beneficiaries per year[b]					
Year 6	1,761	2,076	-315	25 (66)	1.4
Year 5	1,765	2,141	-376	-36 (60)	-2.1
Year 4	1,631	1,950	-319	21 (74)	1.2
Year 3	1,738	2,066	-328	12 (66)	0.7
Year 2	1,761	2,064	-303	37 (57)	2.2
Year 1	1,793	2,085	-292	48 (39)	2.8
One year pre-IAH[c]	1,850	2,190	-340	-	-
Two years pre-IAH	1,945	2,288	-344	-4 (31)	-0.2
Number of hospital admissions preceded by an ED visit per 1,000 beneficiaries per year					
Year 6	1,421	1,598	-178	-33 (61)	-2.3
Year 5	1,454	1,679	-225	-81 (73)	-5.7
Year 4	1,318	1,513	-195	-51 (72)	-3.6
Year 3	1,421	1,615	-194	-49 (60)	-3.5
Year 2	1,437	1,589	-153	-8 (57)	-0.6
Year 1	1,489	1,639	-150	-6 (39)	-0.4
One year pre-IAH[c]	1,567	1,712	-144	-	-
Two years pre-IAH	1,668	1,805	-137	8 (29)	0.6
Number of outpatient ED visits per 1,000 beneficiaries per year[d]					
Year 6	1,755	1,797	-42	-25 (102)	-1.8
Year 5	1,671	1,717	-46	-29 (100)	-2.0
Year 4	1,613	1,642	-30	-13 (78)	-0.9
Year 3	1,695	1,698	-3	13 (71)	0.9
Year 2	1,612	1,542	70	86 (56)	6.0
Year 1	1,440	1,459	-19	-2 (57)	-0.2
One year pre-IAH[c]	1,467	1,484	-16	-	-
Two years pre-IAH	1,436	1,464	-28	-12 (44)	-0.8

Source: Mathematica's analysis of data from the IAH implementation contractor and Medicare claims and enrollment data from the Chronic Conditions Warehouse.

Notes: Total unweighted number of observations across all years is 290,514. Standard errors are in parentheses. We computed coefficients and standard errors by using the weighted sample size, which considers the matching and eligibility weights. The table reports the regression-adjusted means of the IAH and matched comparison groups in each year. We obtained regression-adjusted means by applying the estimated regression coefficients to the covariates of IAH beneficiaries in the latest demonstration year. Because of

Exhibit B.11b (*continued*)

rounding, a difference-in-differences estimate displayed as zero may be shown alongside a percentage effect that exceeds zero.

*/**/*** The difference is statistically significant at the 0.10/0.05/0.01 level.

[a] We used the unadjusted IAH group mean in the year before the demonstration to calculate the percentage effect for each demonstration year. Exhibit B.1 reports the baseline unadjusted IAH group mean for all outcomes.

[b] The number of hospital admissions includes observation stays.

[c] The difference-in-differences estimate for the year before the demonstration is zero (with no standard error) in all regressions because we calculated that estimate for each year as the difference in means between IAH and comparison beneficiaries in that year minus the difference in the year before the demonstration.

[d] The number of outpatient ED visits reflects all those not resulting in a hospital admission, including those resulting in an observation stay.

ED = emergency department.

Exhibit B.12. Estimated effect of IAH on hospital care use in Years 1 to 7, by number of ADLs with which beneficiary required assistance from another person, practices that participated in Year 7

Service type	IAH	Comparison	Difference (IAH - comparison)	Difference-in-differences estimated effect	Percentage effect[a]
5-6 ADL, number of hospital admissions per 1,000 beneficiaries per year[b]					
Year 7	1800	2269	-469	-130 (83)	-6.4
Year 6	2090	2418	-328	11 (96)	0.5
Year 5	2118	2509	-391	-52 (83)	-2.6
Year 4	1963	2325	-362	-23 (82)	-1.1
Year 3	2092	2396	-304	34 (88)	1.7
Year 2	2099	2362	-263	76 (58)	3.7
Year 1	2137	2437	-299	39 (81)	1.9
One year pre-IAH[c]	2160	2499	-339	-	-
Two years pre-IAH	2284	2654	-370	-31 (54)	-1.5
5-6 ADL, number of hospital admissions preceded by an ED visit per 1,000 beneficiaries per year					
Year 7	1518	1869	-351	-208 (77)***	-12.1
Year 6	1728	1909	-182	-39 (82)	-2.2
Year 5	1790	2012	-222	-79 (70)	-4.6
Year 4	1647	1862	-215	-72 (68)	-4.2
Year 3	1754	1922	-168	-25 (71)	-1.4
Year 2	1744	1850	-106	37 (51)	2.2
Year 1	1813	1953	-140	3 (71)	0.2
One year pre-IAH[c]	1846	1989	-143	-	-
Two years pre-IAH	1960	2118	-157	-14 (54)	-0.8
5-6 ADL, number of outpatient ED visits per 1,000 beneficiaries per year[d]					
Year 7	1375	1521	-147	-84 (113)	-6.4
Year 6	1664	1825	-161	-98 (125)	-7.5
Year 5	1623	1762	-139	-76 (105)	-5.8
Year 4	1549	1638	-89	-26 (102)	-2.0
Year 3	1586	1743	-157	-94 (87)	-7.3
Year 2	1621	1553	67	130 (82)	10.0
Year 1	1490	1478	11	74 (80)	5.7
One year pre-IAH[c]	1421	1484	-63	-	-
Two years pre-IAH	1454	1459	-5	58 (74)	4.5

Source: Mathematica's analysis of data from the IAH implementation contractor and Medicare claims and enrollment data from the Chronic Conditions Warehouse.

Notes: Subgroup estimates are calculated using one regression that interacts a binary indicator for subgroup with all covariates in the model, once for total spending and once for inpatient spending. A Wald test was used

Exhibit B.12 (*continued*)

to compare Year 7 effect estimates in each subgroup. Standard errors are in parentheses. We computed coefficients and standard errors by using the weighted sample size, which considers the matching and eligibility weights. The table reports the regression-adjusted means of the IAH and matched comparison groups in each year. We obtained regression-adjusted means by applying the estimated regression coefficients to the covariates of IAH beneficiaries in the subgroup in the latest demonstration year. Because of rounding, a difference-in-differences estimate displayed as zero may be shown alongside a percentage effect that exceeds zero. Estimated effects in Year 7 are interpreted as effects of IAH during the first year of the COVID-19 pandemic.

*/**/*** The difference is statistically significant at the 0.10/0.05/0.01 level.

[a] We used the unadjusted IAH subgroup mean in the year before the demonstration to calculate the percentage effect for each demonstration year.

[b] The number of hospital admissions includes observation stays.

[c] The difference-in-differences estimate for the period before the demonstration is zero (with no standard error) in all regressions because we calculated that estimate for each year as the difference in means between IAH and comparison beneficiaries in that year minus the difference in the year before the demonstration.

[d] The number of outpatient ED visits reflects all those not resulting in a hospital admission, including those resulting in an observation stay.

ADL = activities of daily living; PBPM = per beneficiary per month.

Exhibit B.13a. Estimated effect of IAH on potentially avoidable hospital admissions, potentially avoidable outpatient ED visits, and probability of unplanned readmission in Years 1 to 7, practices that participated in Year 7

Outcome	IAH	Comparison	Difference (IAH - comparison)	Difference-in-differences estimated effect	Percentage effect[a]
Number of potentially avoidable hospital admissions per 1,000 beneficiaries per year[b]					
Year 7	306	395	-89	-27 (31)	-5.9
Year 6	416	476	-60	2 (32)	0.5
Year 5	492	577	-85	-23 (27)	-5.1
Year 4	389	485	-96	-34 (30)	-7.6
Year 3	416	505	-89	-27 (25)	-6.0
Year 2	469	495	-26	36 (23)	8.0
Year 1	480	531	-50	12 (17)	2.6
One year pre-IAH[c]	498	560	-62	-	-
Two years pre-IAH	529	605	-76	-14 (17)	-3.1
Number of potentially avoidable outpatient ED visits per 1,000 beneficiaries per year[d]					
Year 7	180	196	-16	3 (15)	1.4
Year 6	254	263	-9	10 (18)	5.3
Year 5	251	255	-5	14 (14)	7.3
Year 4	214	248	-34	-15 (12)	-8.2
Year 3	228	240	-12	7 (21)	3.5
Year 2	218	220	-2	16 (16)	8.6
Year 1	205	205	1	19 (13)	10.3
One year pre-IAH[c]	202	220	-18	-	-
Two years pre-IAH	214	212	2	21 (16)	11.0
Probability (as a percentage) of having a qualifying hospital discharge and unplanned readmission within 30 days of discharge					
Year 7	15.74	18.60	-2.86	-0.55 (0.99)	-3.4
Year 6	17.70	19.43	-1.73	0.57 (1.16)	3.5
Year 5	18.01	19.97	-1.95	0.35 (1.22)	2.2
Year 4	15.31	17.49	-2.18	0.13 (1.31)	0.8
Year 3	17.80	19.97	-2.17	0.13 (1.00)	0.8
Year 2	18.54	20.00	-1.46	0.84 (1.15)	5.2

Exhibit B.13a (*continued*)

Outcome	IAH	Comparison	Difference (IAH - comparison)	Difference-in-differences estimated effect	Percentage effect[a]
Year 1	19.25	19.84	-3.00	1.71* (0.89)	10.5
One year pre-IAH[c]	18.74	21.04	-2.30	-	-
Two years pre-IAH	21.43	22.95	-1.52	0.78 (0.77)	4.8

Source: Mathematica's analysis of data from the IAH implementation contractor and Medicare claims and enrollment data from the Chronic Conditions Warehouse.

Notes: Total unweighted number of observations across all years is 259,609. We define a potentially avoidable hospital admission (or outpatient ED visit) as one in which appropriate primary and specialty care may prevent or reduce the need for a hospital admission (or ED visit). Standard errors are in parentheses. We computed coefficients and standard errors by using the weighted sample size, which considers the matching and eligibility weights. The table reports the regression-adjusted means of the IAH and matched comparison groups in each year. We obtained regression-adjusted means by applying the estimated regression coefficients to the covariates of IAH beneficiaries in the latest demonstration year. Because of rounding, a difference-in-differences estimate displayed as zero may be shown alongside a percentage effect that exceeds zero. Estimated effects in Year 7 are interpreted as effects of IAH during the first year of the COVID-19 pandemic.

*/**/*** The difference is statistically significant at the 0.10/0.05/0.01 level.

[a] We used the unadjusted IAH group mean in the year before the demonstration to calculate the percentage effect for each demonstration year. Exhibit B.1 reports the baseline unadjusted IAH group mean for all outcomes.

[b] The number of hospital admissions includes observation stays.

[c] The difference-in-differences estimate for the period before the demonstration is zero (with no standard error) in all regressions because we calculated that estimate for each year as the difference in means between IAH and comparison beneficiaries in that year minus the difference in the year before the demonstration.

[d] The number of outpatient ED visits reflects those not resulting in hospital admission, including those resulting in an observation stay. The measure excluded ED visits that led to a hospital admission because there was no diagnosis from such a visit in a claim record when it led to a hospital admission.

ED = emergency department.

Exhibit B.13b. Estimated effect of the IAH payment incentive on potentially avoidable hospital admissions, potentially avoidable outpatient ED visits, and probability of unplanned readmission in Years 1 to 6, practices that participated in Year 6

Outcome	IAH	Comparison	Difference (IAH - comparison)	Difference-in-differences estimated effect	Percentage effect[a]
Number of potentially avoidable hospital admissions per beneficiary per year[b]					
Year 6	385	454	-69	-8 (27)	-1.9
Year 5	454	551	-97	-36 (24)	-8.3
Year 4	366	463	-98	-37 (25)	-8.4
Year 3	397	488	-92	-31 (21)	-7.0
Year 2	430	479	-49	12 (22)	2.6
Year 1	450	509	-58	3 (14)	0.6
One year pre-IAH[c]	472	533	-61	-	-
Two years pre-IAH	500	580	-80	-19 (15)	-4.3
Number of potentially avoidable outpatient ED visits per beneficiary per year[d]					
Year 6	254	254	0	17 (17)	9.0
Year 5	244	260	-16	1 (12)	0.5
Year 4	215	246	-31	-14 (13)	-7.5
Year 3	220	230	-10	6 (17)	3.5
Year 2	211	215	-3	13 (13)	7.1
Year 1	193	200	-8	9 (13)	5.0
One year pre-IAH[c]	194	211	-17	-	-
Two years pre-IAH	200	202	-2	15 (14)	8.0
Probability (as a percentage) of having a qualifying hospital discharge and unplanned readmission within 30 days of discharge					
Year 6	15.91	18.17	-2.26	-0.02 (1.05)	-0.1
Year 5	15.94	18.56	-2.62	-0.38 (1.20)	-2.4
Year 4	13.71	16.23	-2.52	-0.28 (1.15)	-1.8
Year 3	16.06	19.06	-3.00	-0.76 (0.94)	-4.9
Year 2	16.67	18.71	-2.04	0.20 (1.09)	1.3
Year 1	17.64	18.74	-1.10	1.13 (0.81)	7.2
One year pre-IAH[c]	17.47	19.70	-2.24	-	-
Two years pre-IAH	20.00	21.90	-1.90	0.34 (0.68)	2.2

Source: Mathematica's analysis of data from the IAH implementation contractor and Medicare claims and enrollment data from the Chronic Conditions Warehouse.

Notes: Total unweighted number of observations across all years is 290,514. We define a potentially avoidable hospital admission (or outpatient ED visit) as one in which appropriate primary and specialty care may prevent or reduce the need for a hospital admission (or ED visit). Standard errors are in parentheses. We

Exhibit B.13b (*continued*)

computed coefficients and standard errors by using the weighted sample size, which considers the matching and eligibility weights. The table reports the regression-adjusted means of the IAH and matched comparison groups in each year. We obtained regression-adjusted means by applying the estimated regression coefficients to the covariates of IAH beneficiaries in the latest demonstration year. Because of rounding, a difference-in-differences estimate displayed as zero may be shown alongside a percentage effect that exceeds zero.

*/**/*** The difference is statistically significant at the 0.10/0.05/0.01 level.

[a] We used the unadjusted IAH group mean in the year before the demonstration to calculate the percentage effect for each demonstration year. Exhibit B.1 reports the baseline unadjusted IAH group mean for all outcomes.

[b] The number of hospital admissions includes observation stays.

[c] The difference-in-differences estimate for the period before the demonstration is zero (with no standard error) in all regressions because we calculated that estimate for each year as the difference in means between IAH and comparison beneficiaries in that year minus the difference in the year before the demonstration.

[d] The number of outpatient ED visits reflects those not resulting in hospital admission, including those resulting in an observation stay. The measure excluded ED visits that led to a hospital admission because there was no diagnosis from such a visit in a claim record when it led to a hospital admission.

ED = emergency department.

Exhibit B.14. Estimated effect of IAH on potentially avoidable hospital admissions and potentially avoidable outpatient ED visits, in Years 1 to 7, by number of ADLs with which beneficiary required assistance from another person, practices that participated in Year 7

Outcome	IAH	Comparison	Difference (IAH - comparison)	Difference-in-differences estimated effect	Percentage effect[a]
5-6 ADL, number of potentially avoidable hospital admissions per 1,000 beneficiaries per year[b]					
Year 7	321	432	-111	-45 (32)	-9.1
Year 6	445	507	-61	5 (39)	0.9
Year 5	522	612	-90	-24 (34)	-4.8
Year 4	421	526	-104	-38 (32)	-7.7
Year 3	471	532	-61	5 (33)	1.1
Year 2	518	527	-10	56 (29)**	11.4
Year 1	541	579	-38	28 (26)	5.6
One year pre-IAH[c]	534	600	-66	-	-
Two years pre-IAH	592	662	-70	-4 (30)	-0.9
5-6 ADL, number of potentially avoidable outpatient ED visits per 1,000 beneficiaries per year[d]					
Year 7	190	210	-19	3 (20)	1.9
Year 6	268	270	-2	21 (21)	11.9
Year 5	267	268	-1	22 (19)	12.3
Year 4	232	251	-19	4 (17)	2.3
Year 3	247	247	0	23 (19)	12.8
Year 2	229	229	0	23 (20)	12.9
Year 1	216	218	-2	21 (20)	11.7
One year pre-IAH[c]	204	227	-23	-	-
Two years pre-IAH	232	219	13	36 (25)	19.8

Source: Mathematica's analysis of data from the IAH implementation contractor and Medicare claims and enrollment data from the Chronic Conditions Warehouse.

Notes: Total unweighted number of observations across all years is 259,609. We define a potentially avoidable hospital admission (or outpatient ED visit) as one in which appropriate primary and specialty care may prevent or reduce the need for a hospital admission (or ED visit). Standard errors are in parentheses. We computed coefficients and standard errors by using the weighted sample size, which considers the matching and eligibility weights. The table reports the regression-adjusted means of the IAH and matched comparison groups in each year. We obtained regression-adjusted means by applying the estimated regression coefficients to the covariates of IAH beneficiaries in the latest demonstration year. Because of rounding, a difference-in-differences estimate displayed as zero may be shown alongside a percentage effect that exceeds zero. Estimated effects in Year 7 are interpreted as effects of IAH during the first year of the COVID-19 pandemic.

*/**/*** The difference is statistically significant at the 0.10/0.05/0.01 level.

Exhibit B.14 (*continued*)

[a] We used the unadjusted IAH subgroup mean in the year before the demonstration to calculate the percentage effect for each demonstration year. Exhibit B.1 reports the baseline unadjusted IAH group mean for all outcomes.

[b] The number of hospital admissions includes observation stays.

[c] The difference-in-differences estimate for the period before the demonstration is zero (with no standard error) in all regressions because we calculated that estimate for each year as the difference in means between IAH and comparison beneficiaries in that year minus the difference in the year before the demonstration.

[d] The number of outpatient ED visits reflects those not resulting in hospital admission, including those resulting in an observation stay. The measure excluded ED visits that led to a hospital admission because there was no diagnosis from such a visit in a claim record when it led to a hospital admission.

ED = emergency department; ADL = activities of daily living.

Exhibit B.15a. Estimated effect of IAH on probability (as a percentage) of dying within the demonstration year in Years 1 to 7, practices that participated in Year 7

	IAH	Comparison	Difference (IAH - comparison)	Difference-in-differences estimated effect	Percentage effect[a]
Year 7	15.83	24.92	-9.10	-2.42*** (0.88)	-16.8
Year 6	15.08	21.02	-5.94	0.73 (1.09)	5.1
Year 5	15.16	21.83	-6.67	0.01 (0.79)	0.1
Year 4	14.99	21.96	-6.97	-0.29 (0.82)	-2.0
Year 3	17.16	21.47	-4.31	2.36*** (0.65)	16.4
Year 2	16.18	21.67	-5.49	1.18 (0.83)	8.2
Year 1	16.76	22.39	-5.63	1.04* (0.56)	7.2
One year pre-IAH[b]	16.44	23.11	-6.67	-	-
Two years pre-IAH	15.53	24.18	-8.64	-1.97** (0.94)	-13.7

Source: Mathematica's analysis of data from the IAH implementation contractor and Medicare claims and enrollment data from the Chronic Conditions Warehouse.

Notes: Total unweighted number of observations across all years is 259,609. Standard errors are in parentheses. We computed coefficients and standard errors by using the weighted sample size, which considers the matching and eligibility weights. The table reports the regression-adjusted means of the IAH and matched comparison groups in each year. Regression-adjusted means are obtained by applying the estimated regression coefficients to the covariates of IAH beneficiaries in the latest demonstration year. Because of rounding, a difference-in-differences estimate displayed as zero may be shown alongside a percentage effect that exceeds zero. Estimated effects in Year 7 are interpreted as effects of IAH during the first year of the COVID-19 pandemic.

*/**/*** The difference is statistically significant at the 0.10/0.05/0.01 level.

[a] We used the unadjusted IAH group mean in the year before the demonstration to calculate the percentage effect for each demonstration year. Exhibit B.1 reports the baseline unadjusted IAH group mean for all outcomes.

[b] The difference-in-differences estimate for the year before the demonstration is zero (with no standard error) in all regressions because we calculated that estimate for each year as the difference in means between IAH and comparison beneficiaries in that year minus the difference in the year before the demonstration.

Exhibit B.15b. Estimated effect of the IAH payment incentive on probability (as a percentage) of dying within the demonstration year in Years 1 to 6, practices that participated in Year 6

	IAH	Comparison	Difference (IAH - comparison)	Difference-in-differences estimated effect	Percentage effect[a]
Year 6	15.17	21.11	-5.94	0.56 (0.86)	3.8
Year 5	14.97	21.92	-6.95	-0.45 (0.69)	-3.1
Year 4	14.82	21.94	-7.11	-0.61 (0.70)	-4.2
Year 3	16.51	21.24	-4.73	1.77** (0.71)	12.1
Year 2	15.74	21.31	-5.57	0.93 (0.75)	6.4
Year 1	16.23	22.01	-5.78	0.72 (0.62)	4.9
One year pre-IAH[b]	16.32	22.82	-6.50	-	-
Two years pre-IAH	15.61	23.75	-8.14	-1.64** (0.79)	-11.2

Source: Mathematica's analysis of data from the IAH implementation contractor and Medicare claims and enrollment data from the Chronic Conditions Warehouse.

Notes: Total unweighted number of observations across all years is 290,514. Standard errors are in parentheses. We computed coefficients and standard errors by using the weighted sample size, which considers the matching and eligibility weights. The table reports the regression-adjusted means of the IAH and matched comparison groups in each year. Regression-adjusted means are obtained by applying the estimated regression coefficients to the covariates of IAH beneficiaries in the latest demonstration year. Because of rounding, a difference-in-differences estimate displayed as zero may be shown alongside a percentage effect that exceeds zero.

*/**/*** The difference is statistically significant at the 0.10/0.05/0.01 level.

[a] We used the unadjusted IAH group mean in the year before the demonstration to calculate the percentage effect for each demonstration year. Exhibit B.1 reports the baseline unadjusted IAH group mean for all outcomes.

[b] The difference-in-differences estimate for the year before the demonstration is zero (with no standard error) in all regressions because we calculated that estimate for each year as the difference in means between IAH and comparison beneficiaries in that year minus the difference in the year before the demonstration.

Exhibit B.16a. Estimated effect of IAH on probability (as a percentage) of dying within the demonstration year in Years 1 to 7, controlling for COVID-19 diagnosis in Year 7

	IAH	Comparison	Difference (IAH - comparison)	Difference-in-differences estimated effect	Percentage effect[a]
Year 7	15.82	24.83	-9.01	-2.24** (0.88)	-15.6
Year 6	15.49	21.52	-6.03	0.74 (1.11)	5.1
Year 5	15.58	22.34	-6.77	-0.00 (0.80)	0.0
Year 4	15.41	22.48	-7.07	-0.30 (0.83)	-2.1
Year 3	17.60	21.98	-4.38	2.39*** (0.66)	16.6
Year 2	16.61	22.17	-5.57	1.20 (0.84)	8.3
Year 1	17.20	22.91	-5.71	1.06* (0.57)	7.3
One year pre-IAH[b]	16.87	23.64	-6.77	-	-
Two years pre-IAH	15.95	24.72	-8.77	-2.00** (0.95)	-13.9

Source: Mathematica's analysis of data from the IAH implementation contractor and Medicare claims and enrollment data from the Chronic Conditions Warehouse.

Notes: Total unweighted number of observations across all years is 259,609. Standard errors are in parentheses. We computed coefficients and standard errors by using the weighted sample size, which considers the matching and eligibility weights. The table reports the regression-adjusted means of the IAH and matched comparison groups in each year. Regression-adjusted means are obtained by applying the estimated regression coefficients to the covariates of IAH beneficiaries in the latest demonstration year. Because of rounding, a difference-in-differences estimate displayed as zero may be shown alongside a percentage effect that exceeds zero. Estimated effects in Year 7 are interpreted as effects of IAH during the first year of the COVID-19 pandemic.

*/**/*** The difference is statistically significant at the 0.10/0.05/0.01 level.

[a] We used the unadjusted IAH group mean in the year before the demonstration to calculate the percentage effect for each demonstration year. Exhibit B.1 reports the baseline unadjusted IAH group mean for all outcomes.

[b] The difference-in-differences estimate for the year before the demonstration is zero (with no standard error) in all regressions because we calculated that estimate for each year as the difference in means between IAH and comparison beneficiaries in that year minus the difference in the year before the demonstration.

Exhibit B.16b. Estimated effect of IAH on probability (as a percentage) of dying within the demonstration year in Years 1 to 7, controlling for COVID-19 hospitalization in Year 7

	IAH	Comparison	Difference (IAH - comparison)	Difference-in-differences estimated effect	Percentage effect[a]
Year 7	15.79	24.26	-8.46	-1.71* (0.88)	-11.9
Year 6	15.59	21.60	-6.01	0.74 (1.11)	5.1
Year 5	15.67	22.43	-6.75	0.00 (0.81)	0.0
Year 4	15.50	22.55	-7.05	-0.30 (0.83)	-2.1
Year 3	17.70	22.06	-4.36	2.39*** (0.67)	16.6
Year 2	16.70	22.25	-5.55	1.20 (0.84)	8.32
Year 1	17.29	22.98	-5.69	1.06* (0.57)	7.36
One year pre-IAH[b]	16.97	23.72	-6.75	-	-
Two years pre-IAH	16.04	24.80	-8.76	-2.01* (0.95)	-13.9

Source: Mathematica's analysis of data from the IAH implementation contractor and Medicare claims and enrollment data from the Chronic Conditions Warehouse.

Notes: Total unweighted number of observations across all years is 259,609. Standard errors are in parentheses. We computed coefficients and standard errors by using the weighted sample size, which considers the matching and eligibility weights. The table reports the regression-adjusted means of the IAH and matched comparison groups in each year. Regression-adjusted means are obtained by applying the estimated regression coefficients to the covariates of IAH beneficiaries in the latest demonstration year. Because of rounding, a difference-in-differences estimate displayed as zero may be shown alongside a percentage effect that exceeds zero. Estimated effects in Year 7 are interpreted as effects of IAH during the first year of the COVID-19 pandemic.

*/**/*** The difference is statistically significant at the 0.10/0.05/0.01 level.

[a] We used the unadjusted IAH group mean in the year before the demonstration to calculate the percentage effect for each demonstration year. Exhibit B.1 reports the baseline unadjusted IAH group mean for all outcomes.

[b] The difference-in-differences estimate for the year before the demonstration is zero (with no standard error) in all regressions because we calculated that estimate for each year as the difference in means between IAH and comparison beneficiaries in that year minus the difference in the year before the demonstration.

Exhibit B.17. Unadjusted mean total Medicare spending PBPM for decedents and non-decedents, by year, practices that participated in Year 7

	Year 6		Year 7		Change from Year 6 to Year 7	
	IAH	Comparison	IAH	Comparison	IAH	Comparison
Decedents	$7,958	$10,459	$8,217	$10,942	3.3%	4.6%
Non-decedents	$4,630	$4,830	$4,529	$4,884	-2.2%	1.1%
All beneficiaries	$5,025	$5,488	$5,002	$5,736	-0.5%	4.5%

Source: Mathematica's analysis of data from the IAH implementation contractor and Medicare claims and enrollment data from the Chronic Conditions Warehouse.

Notes: Results are not regression-adjusted.

PBPM = per beneficiary per month.

Exhibit B.18a. Estimated effect of IAH on probability (as a percentage) of entering institutional long-term care within the demonstration year in Years 1 to 7, practices that participated in Year 7

	IAH	Comparison	Difference (IAH - comparison)	Difference-in-differences estimated effect	Percentage effect[a]
Year 7	5.50	13.13	-7.63***	-0.07 (0.62)	-0.9
Year 6	6.50	12.14	-5.64***	1.92*** (0.66)	27.3
Year 5	7.52	13.34	-5.82***	1.75** (0.75)	24.7
Year 4	7.90	14.98	-7.08***	0.48 (0.66)	6.8
Year 3	8.81	15.42	-6.61***	0.95 (0.67)	13.5
Year 2	8.41	14.41	-6.00***	1.56*** (0.60)	22.1
Year 1	9.23	15.07	-5.84***	1.72*** (0.54)	24.4
One year pre-IAH[b]	8.97	16.54	-7.56***	-	-
Two years pre-IAH	8.91	17.73	-8.82***	-1.26* (0.66)	-17.9

Source: Mathematica's analysis of data from the IAH implementation contractor and Medicare claims and enrollment data from the Chronic Conditions Warehouse.

Notes: Total unweighted number of observations across all years is 259,609. Standard errors are in parentheses. We computed coefficients and standard errors by using the weighted sample size, which considers the matching and eligibility weights. The table reports the regression-adjusted means of the IAH and matched comparison groups in each year. We obtained regression-adjusted means by applying the estimated regression coefficients to the covariates of IAH beneficiaries in the latest demonstration year. Because of rounding, a difference-in-differences estimate displayed as zero may be shown alongside a percentage effect that exceeds zero. Estimated effects in Year 7 are interpreted as effects of IAH during the first year of the COVID-19 pandemic.

*/**/*** The difference is statistically significant at the 0.10/0.05/0.01 level.

[a] We used the unadjusted IAH group mean in the year before the demonstration to calculate the percentage effect for each demonstration year. Exhibit B.1 reports the baseline unadjusted IAH group mean for all outcomes.

[b] The difference-in-differences estimate for the year before the demonstration is zero (with no standard error) in all regressions because we calculated that estimate for each year as the difference in means between IAH and comparison beneficiaries in that year minus the difference in the year before the demonstration.

Exhibit B.18b. Estimated effect of the IAH payment incentive on probability (as a percentage) of entering institutional long-term care within the demonstration year in Years 1 to 6, practices that participated in Year 6

	IAH	Comparison	Difference (IAH - comparison)	Difference-in-differences estimated effect	Percentage effect[a]
Year 6	6.58	12.02	-5.44	2.38*** (0.60)	34.4
Year 5	7.02	13.57	-6.55	1.27* (0.74)	18.4
Year 4	7.59	14.89	-7.30	0.51 (0.60)	7.4
Year 3	8.42	15.62	-7.20	0.62 (0.62)	8.9
Year 2	8.24	15.03	-6.79	1.03* (0.56)	14.9
Year 1	8.71	15.44	-6.74	1.08** (0.54)	15.6
One year pre-IAH[b]	8.96	16.78	-7.82	-	-
Two years pre-IAH	8.99	17.94	-8.95	-1.13* (0.59)	-16.4

Source: Mathematica's analysis of data from the IAH implementation contractor and Medicare claims and enrollment data from the Chronic Conditions Warehouse.

Notes: Total unweighted number of observations across all years is 290,514. Standard errors are in parentheses. We computed coefficients and standard errors by using the weighted sample size, which considers the matching and eligibility weights. The table reports the regression-adjusted means of the IAH and matched comparison groups in each year. We obtained regression-adjusted means by applying the estimated regression coefficients to the covariates of IAH beneficiaries in the latest demonstration year. Because of rounding, a difference-in-differences estimate displayed as zero may be shown alongside a percentage effect that exceeds zero.

*/**/*** The difference is statistically significant at the 0.10/0.05/0.01 level.

[a] We used the unadjusted IAH group mean in the year before the demonstration to calculate the percentage effect for each demonstration year. Exhibit B.1 reports the baseline unadjusted IAH group mean for all outcomes.

[b] The difference-in-differences estimate for the year before the demonstration is zero (with no standard error) in all regressions because we calculated that estimate for each year as the difference in means between IAH and comparison beneficiaries in that year minus the difference in the year before the demonstration.

References

Bosworth, A., K. Finegold, L.W. Samson, S. Sheingold, W. Tarazi, and R. Zuckerman. "Risk of COVID-19 Infection, Hospitalization, and Death in Fee-For-Service Medicare." Washington, DC: Office of the Assistant Secretary for Planning and Evaluation, U.S. Department of Health and Human Services, 2021. Available at https://aspe.hhs.gov/sites/default/files/private/pdf/265271/risk-score-issue-brief.pdf. Accessed November 11, 2021.

Centers for Medicare & Medicaid Services (CMS). "Chronic Conditions Among Medicare Beneficiaries, CMS Chartbook: 2012 Edition." Baltimore, MD: CMS, 2012. Available at https://www.cms.gov/research-statistics-data-and-systems/statistics-trends-and-reports/chronic-conditions/downloads/2012chartbook.pdf.

Centers for Medicare & Medicaid Services (CMS). "Risk Adjustment." Baltimore, MD: CMS, 2019. Available at https://www.cms.gov/Medicare/Health-Plans/MedicareAdvtgSpecRateStats/Risk-Adjustors.html.

Centers for Medicare & Medicaid Services (CMS). "Shared Savings Program Fast Facts – As of January 1, 2021." Baltimore, MD: CMS, 2021a. Available at https://www.cms.gov/files/document/2021-shared-savings-program-fast-facts.pdf. Accessed June 9, 2022.

Centers for Medicare & Medicaid Services (CMS). "Independence at Home Demonstration Revised Actuarial Shared Savings Methodology." Baltimore, MD: CMS, 2021b. Available at https://innovation.cms.gov/media/document/iah-rev-act-meth-specs. Accessed May 26, 2022.

Centers for Medicare & Medicaid Services (CMS). "Independence at Home Demonstration Performance Year 7 Results." Baltimore, MD: CMS, 2022. Available at https://innovation.cms.gov/media/document/iah-py7-fs. Accessed May 26, 2022.

Damschroder L., D. Aron, R. Keith, S. Kirsh, J. Alexander, and J. Lowery. "Fostering Implementation of Health Services Research Findings into Practice: A Consolidated Framework for Advancing Implementation Science." *Implementation Science,* vol. 4, no. 50, 2009.

Deb, P. "Estimating the Effect of Treatment on Healthcare Cost in the Presence of Death." Unpublished manuscript, March 2016.

Gagne, J.J., R.J. Glynn, J. Avorn, R. Levin, and S. Schneeweiss. "A Combined Comorbidity Score Predicted Mortality in Elderly Patients Better than Existing Scores." *Journal of Clinical Epidemiology,* vol. 64, no. 7, 2011, pp. 749–759.

Kaiser Family Foundation. "Dual Eligibles as a Percent of Total Medicare Beneficiaries." Available at https://www.kff.org/medicaid/state-indicator/duals-as-a-of-medicare-beneficiaries/. Accessed March 24, 2021.

Kandilov, A., M. Ingber, M. Morley, N. Coomer, K. Dalton, B. Gage, C. Superina, and D. Kennell. "Chronically Critically Ill Population Payment Recommendations (CCIP-PR)." Research Triangle Park, NC: RTI International, 2014. Available at https://innovation.cms.gov/files/reports/chronicallycriticallyillpopulation-report.pdf.

Lewandowski, D., D. Kurowicka, and H. Joe. "Generating Random Correlation Matrices Based on Vines and Extended Onion Method." *Journal of Multivariate Analysis,* vol. 100, 2009, pp. 1989–2001.

Navathe, A.S., C.T. Dinh, A. Chen, and J.M. Liao. "Findings and Implications from MIPS Year 1 Performance Data." Health Affairs blog, January 18, 2019. Available at https://www.healthaffairs.org/do/10.1377/hblog20190117.305369/full/. Accessed June 9, 2021.

Rubin, D.B. "Using Propensity Scores to Help Design Observational Studies: Application to the Tobacco Litigation." *Health Services and Outcomes Research Methodology,* vol. 2, no. 3–4, 2001, pp. 169–188.

Stuart, E.A. "Matching Methods for Causal Inference: A Review and Look Forward." *Statistical Science,* vol. 21, no. 1, 2010, pp. 1–25.

Mathematica Inc.

Princeton, NJ • Ann Arbor, MI • Cambridge, MA
Chicago, IL • Oakland, CA • Seattle, WA
Tucson, AZ • Woodlawn, MD • Washington, DC

Mathematica®
Progress Together

mathematica.org website

EDI Global, a Mathematica Company

Operating in Tanzania, Uganda, Kenya, and the United Kingdom