Evaluation of the Independence at Home Demonstration Final Report—Appendices

August 2026

Laura Kimmey, Jason Rotter, Lauren Forrow, Joey Lovins, Rumin Sarwar

Contributing staff: Megan Baker, Linda Barterian, Christine Cheu, Irna May Connor, Tyler Fisher, Emily Hernandez, Ian Huff, David Jones, Joy Nyabwari, Maya Palakal, Tyler Rose, Carol Urato

Submitted to:

Centers for Medicare & Medicaid Services
Center for Medicare and Medicaid Innovation
7500 Security Blvd.
Baltimore, MD 21244
Project Officer: David Nyweide
Contract Number: 75FCMC19D0091/75FCMC23F0002

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

Please download the associated PDFs to your device to ensure accessibility when reading with assistive technology devices.

Contents

Exhibits

Appendix A.

IAH Legislation

The Patient Protection and Affordable Care Act
PUBLIC LAW 111–148—MAR. 23, 2010

SEC. 3024. INDEPENDENCE AT HOME DEMONSTRATION PROGRAM.

Title XVIII of the Social Security Act is amended by inserting after section 1866D, as inserted by section 3023, the following new section:

"INDEPENDENCE AT HOME MEDICAL PRACTICE DEMONSTRATION PROGRAM

"Sec. 1866D. (a) Establishment.—

"(1) IN GENERAL.—The Secretary shall conduct a demonstration program (in this section referred to as the 'demonstration program') to test a payment incentive and service delivery model that utilizes physician and nurse practitioner directed home-based primary care teams designed to reduce expenditures and improve health outcomes in the provision of items and services under this title to applicable beneficiaries (as defined in subsection (d)).

"(2) REQUIREMENT.—The demonstration program shall test whether a model described in paragraph (1), which is accountable for providing comprehensive, coordinated, continuous, and accessible care to high-need populations at home and coordinating health care across all treatment settings, results in—

"(A) reducing preventable hospitalizations;

"(B) preventing hospital readmissions;

"(C) reducing emergency room visits;

"(D) improving health outcomes commensurate with the beneficiaries' stage of chronic illness;

"(E) improving the efficiency of care, such as by reducing duplicative diagnostic and laboratory tests;

"(F) reducing the cost of health care services covered under this title; and

"(G) achieving beneficiary and family caregiver satisfaction.

"(b) Independence at Home Medical Practice.—

"(1) Independence at home medical practice defined.—In this section:

"(A) IN GENERAL.–The term 'independence at home medical practice' means a legal entity that—

"(i) is comprised of an individual physician or nurse practitioner or group of physicians and nurse practitioners that provides care as part of a team that includes physicians, nurses, physician assistants, pharmacists, and other health and social services staff as appropriate who have experience providing home-based primary care to applicable beneficiaries, make in-home visits, and are available 24 hours per day, 7 days per week to carry out plans of care that are tailored to the individual beneficiary's chronic conditions and designed to achieve the results in subsection (a);

"(ii) is organized at least in part for the purpose of providing physicians' services;

"(iii) has documented experience in providing home-based primary care services to high-cost chronically ill beneficiaries, as determined appropriate by the Secretary;

"(iv) furnishes services to at least 200 applicable beneficiaries (as defined in subsection (d)) during each year of the demonstration program;

"(v) has entered into an agreement with the Secretary;

"(vi) uses electronic health information systems, remote monitoring, and mobile diagnostic technology; and

"(vii) meets such other criteria as the Secretary determines to be appropriate to participate in the demonstration program.

The entity shall report on quality measures (in such form, manner, and frequency as specified by the Secretary, which may be for the group, for providers of services and suppliers, or both) and report to the Secretary (in a form, manner, and frequency as specified by the Secretary) such data as the Secretary determines appropriate to monitor and evaluate the demonstration program.

"(B) PHYSICIAN.—The term 'physician' includes, except as the Secretary may otherwise provide, any individual who furnishes services for which payment may be made as physicians' services and has the medical training or experience to fulfill the physician's role described in subparagraph (A)(i).

"(2) PARTICIPATION OF NURSE PRACTITIONERS AND PHYSICIAN ASSISTANTS.—Nothing in this section shall be construed to prevent a nurse practitioner or physician assistant from participating in, or leading, a home-based primary care team as part of an independence at home medical practice if--

"(A) all the requirements of this section are met;

"(B) the nurse practitioner or physician assistant, as the case may be, is acting consistent with State law; and

"(C) the nurse practitioner or physician assistant has the medical training or experience to fulfill the nurse practitioner or physician assistant role described in paragraph (1)(A)(i).

"(3) INCLUSION OF PROVIDERS AND PRACTITIONERS.—Nothing in this subsection shall be construed as preventing an independence at home medical practice from including a provider of services or a participating practitioner described in section 1842(b)(18)(C) that is affiliated with the practice under an arrangement structured so that such provider of services or practitioner participates in the demonstration program and shares in any savings under the demonstration program.

"(4) QUALITY AND PERFORMANCE STANDARDS.—The Secretary shall develop quality performance standards for independence at home medical practices participating in the demonstration program.

"(c) Payment Methodology.—

"(1) ESTABLISHMENT OF TARGET SPENDING LEVEL.—The Secretary shall establish an estimated annual spending target, for the amount the Secretary estimates would have been spent in the absence of the demonstration, for items and services covered under parts A and B furnished to applicable beneficiaries for each qualifying independence at home medical practice under this section. Such spending targets shall be determined on a per capita basis. Such spending targets shall include a risk corridor that takes into account normal variation in expenditures for items and services covered under parts A and B furnished to such beneficiaries with the size of the corridor being related to the number of applicable beneficiaries furnished services by each independence at home medical practice. The spending targets may also be adjusted for other factors as the Secretary determines appropriate.

"(2) INCENTIVE PAYMENTS.—Subject to performance on quality measures, a qualifying independence at home medical practice is eligible to receive an incentive payment under this section if actual expenditures for a year for the applicable beneficiaries it enrolls are less than the estimated spending target established under paragraph (1) for such year. An incentive payment for such year shall be equal to a portion (as determined by the Secretary) of the amount by which actual

expenditures (including incentive payments under this paragraph) for applicable beneficiaries under parts A and B for such year are estimated to be less than 5 percent less than the estimated spending target for such year, as determined under paragraph (1).

"(d) Applicable Beneficiaries.—

"(1) DEFINITION.—In this section, the term 'applicable beneficiary' means, with respect to a qualifying independence at home medical practice, an individual who the practice has determined—

"(A) is entitled to benefits under part A and enrolled for benefits under part B;

"(B) is not enrolled in a Medicare Advantage plan under part C or a PACE program under section 1894;

"(C) has 2 or more chronic illnesses, such as congestive heart failure, diabetes, other dementias designated by the Secretary, chronic obstructive pulmonary disease, ischemic heart disease, stroke, Alzheimer's Disease and neurodegenerative diseases, and other diseases and conditions designated by the Secretary which result in high costs under this title;

"(D) within the past 12 months has had a nonelective hospital admission;

"(E) within the past 12 months has received acute or subacute rehabilitation services;

"(F) has 2 or more functional dependencies requiring the assistance of another person (such as bathing, dressing, toileting, walking, or feeding); and

"(G) meets such other criteria as the Secretary determines appropriate.

"(2) PATIENT ELECTION TO PARTICIPATE.—The Secretary shall determine an appropriate method of ensuring that applicable beneficiaries have agreed to enroll in an independence at home medical practice under the demonstration program. Enrollment in the demonstration program shall be voluntary.

"(3) BENEFICIARY ACCESS TO SERVICES.—Nothing in this section shall be construed as encouraging physicians or nurse practitioners to limit applicable beneficiary access to services covered under this title and applicable beneficiaries shall not be required to relinquish access to any benefit under this title as a condition of receiving services from an independence at home medical practice.

"(e) Implementation.—

"(1) STARTING DATE.—The demonstration program shall begin no later than January 1, 2012. An agreement with an independence at home medical practice under the demonstration program may cover not more than a 3-year period.

"(2) No physician duplication in demonstration participation.—The Secretary shall not pay an independence at home medical practice under this section that participates in section 1899.

"(3) NO BENEFICIARY DUPLICATION IN DEMONSTRATION PARTICIPATION.—The Secretary shall ensure that no applicable beneficiary enrolled in an independence at home medical practice under this section is participating in the programs under section 1899.

"(4) PREFERENCE.—In approving an independence at home medical practice, the Secretary shall give preference to practices that are—

"(A) located in high-cost areas of the country;

"(B) have experience in furnishing health care services to applicable beneficiaries in the home; and

"(C) use electronic medical records, health information technology, and individualized plans of care.

"(5) LIMITATION ON NUMBER OF PRACTICES.—In selecting qualified independence at home medical practices to participate under the demonstration program, the Secretary shall limit the number of such practices so that the number of applicable beneficiaries that may participate in the demonstration program does not exceed 10,000.

"(6) WAIVER.—The Secretary may waive such provisions of this title and title XI as the Secretary determines necessary in order to implement the demonstration program.

"(7) ADMINISTRATION.—Chapter 35 of title 44, United States Code, shall not apply to this section.

"(f) Evaluation and Monitoring.—

"(1) IN GENERAL.—The Secretary shall evaluate each independence at home medical practice under the demonstration program to assess whether the practice achieved the results described in subsection (a).

"(2) MONITORING APPLICABLE BENEFICIARIES.—The Secretary may monitor data on expenditures and quality of services under this title after an applicable beneficiary discontinues receiving services under this title through a qualifying independence at home medical practice.

"(g) REPORTS TO CONGRESS.—The Secretary shall conduct an independent evaluation of the demonstration program and submit to Congress a final report, including best practices under the demonstration program. Such report shall include an analysis of the demonstration program on coordination of care, expenditures under this title, applicable beneficiary access to services, and the quality of health care services provided to applicable beneficiaries.

"(h) FUNDING.—For purposes of administering and carrying out the demonstration program, other than for payments for items and services furnished under this title and incentive payments under subsection (c), in addition to funds otherwise appropriated, there shall be transferred to the Secretary for the Center for Medicare & Medicaid Services Program Management Account from the Federal Hospital Insurance Trust Fund under section 1817 and the Federal Supplementary Medical Insurance Trust Fund under section 1841 (in proportions determined appropriate by the Secretary) $5,000,000 for each of fiscal years 2010 through 2015. Amounts transferred under this subsection for a fiscal year shall be available until expended.

"(i) TERMINATION.—

"(1) MANDATORY TERMINATION.—The Secretary shall terminate an agreement with an independence at home medical practice if—

"(A) the Secretary estimates or determines that such practice will not receive an incentive payment for the second of 2 consecutive years under the demonstration program; or

"(B) such practice fails to meet quality standards during any year of the demonstration program.

"(2) PERMISSIVE TERMINATION.—The Secretary may terminate an agreement with an independence at home medical practice for such other reasons determined appropriate by the Secretary.".

Appendix B.

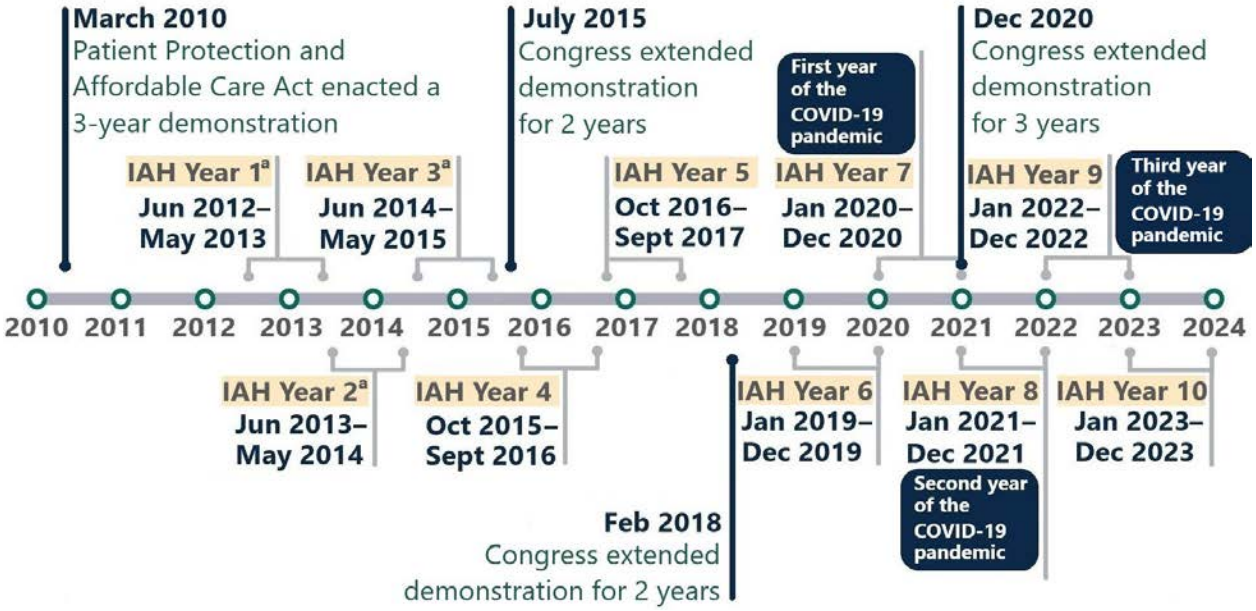
Description of IAH Practices and Beneficiaries

1. Overview

The Independence at Home (IAH) demonstration began in 2012. It was originally intended to last three years, but Congress extended it three times (Exhibit B.1).

This appendix describes the IAH practices and IAH beneficiaries. For IAH practices, we first list the eligibility requirements for IAH practices and summarize when and why practices withdrew from participation. We then describe the practices that participated, including summaries of how the practices operated and provided home-based care during the demonstration period. For IAH beneficiaries, we describe the IAH eligibility requirements for beneficiaries and summarize characteristics of IAH beneficiaries and how they differ from the average Medicare beneficiary. Finally, we list the proposed hypotheses for how home-based primary care (HBPC) and the IAH incentive payment may affect intended outcomes.

Exhibit B.1. Key dates related to the IAH demonstration



[a] For three practices, Years 1 to 3 began in September and ended in August rather than June to May.

2. Description of IAH practices

The law that enacted the IAH demonstration described the eligibility requirements for practices and beneficiaries. Demonstration practices must have prior experience delivering HBPC. Their care teams must also be led by physicians or nurse practitioners (NPs), although care teams could include physician assistants (PAs), clinical staff, and other health and social services staff (Exhibit B.2).

Exhibit B.2. Eligibility requirements for practices to participate in the IAH demonstration

- Be led by physicians or NPs who provide HBPC as part of a team
- Be organized at least in part for the purpose of providing physicians' services
- Have experience providing HBPC to beneficiaries with several chronic illnesses
- Make in-home primary care visits and have staff who are available at all hours
- Use electronic health records, remote monitoring, and mobile diagnostic technology
- Provide services to at least 200 IAH-eligible beneficiaries each year
- Report data for quality measures to the Centers for Medicare & Medicaid Services (CMS)
- Achieve savings at least once in three consecutive years

Of the 18 practices that participated in the IAH demonstration, all but one withdrew before the demonstration ended. Exhibit B.3 summarizes the timing and reasons for their withdrawals and indicates whether the practice was included in the evaluation. Four practices withdrew before Year 4 and were not included in the evaluation. The remaining practices withdrew after Year 4 and were included in the evaluation. The most common reasons for withdrawing included choosing to focus on other priorities (6 practices), internal business issues and reporting difficulties (3 practices), not meeting statutory requirements for participation (3 practices), or beginning a different CMS Innovation Center initiative (3 practices). For the remaining two practices, one withdrew because it did not meet Medicare programmatic billing rules, and another withdrew due to dissatisfaction with demonstration operations.

Exhibit B.3. Reasons why practices left the IAH demonstration

Practice	Timing of withdrawal	Reason for withdrawal	Years included in evaluation
Atlanta, GA (2 organizations)[a]	Before Year 3	Internal business issues and reporting difficulties	None
Chicago, IL (7 organizations)[a]	Before Year 3	Internal business issues and reporting difficulties	None
Stuart, FL (2 organizations)[a]	Before Year 3	Internal business issues and reporting difficulties	None
Louisville, KY	Before Year 4	Did not meet Medicare programmatic billing rules	None
Austin, TX	Before Year 6	Did not meet statutory requirement for achieving savings at least once in three consecutive years	Years 1–5
Cleveland, OH	Before Year 6	Did not meet statutory requirement for achieving savings at least once in three consecutive years	Years 1–5
Boston, MA	Before Year 7	Did not meet statutory requirement for minimum number of beneficiaries	Years 1–6
Durham, NC	Before Year 7	Dissatisfaction with demonstration operations	Years 1–6[b]
Brooklyn, NY	Before Year 8	Began participating in a different CMS Innovation Center initiative	Years 1–7
Portland, OR	Before Year 8	Began participating in a different CMS Innovation Center initiative	Years 1–7
Wilmington, DE	Before Year 8	Began participating in a different CMS Innovation Center initiative	Years 1–7
Dallas, TX	Before Year 10	Chose to focus on other priorities	Years 1–9
Flint, MI	Before Year 10	Chose to focus on other priorities	Years 1–9
Jacksonville, FL	Before Year 10	Chose to focus on other priorities	Years 1–9
Lansing, MI	Before Year 10	Chose to focus on other priorities	Years 1–9
Milwaukee, WI	Before Year 10	Chose to focus on other priorities	Years 1–9
Richmond, VA (3 organizations)[a]	Before Year 10	Chose to focus on other priorities	Years 1–9
Long Island, NY	n.a.	n.a.	Years 1–10

Source: Data from the IAH implementation contractor.

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

[b] Durham left the demonstration late in Year 6, so the practice was not included in the implementation contractor's calculation of quality measure performance and incentive payment results for Year 6. Because Durham participated in most of Year 6, we included the practice in the evaluation in that year.

n.a. = not applicable.

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. Each IAH practice had substantial experience providing HBPC before the demonstration. However, the practices differed substantially in their approaches to care, such as who was included on the care team; whether they were notified automatically of beneficiaries' 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. Below, we summarize care delivery patterns according to each of the three types of practices that participated in the demonstration: HarmonyCares Medical Group (formerly known as Visiting Physicians Association before 2022) practices, academic medical center practices, and independent practices. Information about the practice setting comes from claims data in Years 2 and 4 of the demonstration. We obtained other information from the IAH practices.[1] Exhibits B.4, B.5, and B.6 provide practice-specific information on structural characteristics, operational characteristics, and care delivery processes.

2.1. HarmonyCares Medical Group practices

The five HarmonyCares practices (Dallas, Flint, Jacksonville, Lansing, and Milwaukee) had similar structural characteristics, operational characteristics, and care delivery processes (Exhibits B.4, B.5, and B.6). This medical group is part of a for-profit corporation with multiple HBPC practices operating in several states; five of those practices participated in the demonstration. In each of these practices, most clinicians were physicians. None of the HarmonyCares practices employed a social worker.

Historically, each practice had a patient care coordinator who was the main point of contact for patients and had access to the HarmonyCares corporate infrastructure for finance, human resources, data analytics, and data support. In Year 6, the HarmonyCares 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.[2] HarmonyCares nurse navigators often visited the homes of their patients, although those visits were not billable. In four of the HarmonyCares practices, 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 HarmonyCares practices conducted visits in other settings such as hospitals or skilled nursing facilities (SNFs).

As of Year 6, the HarmonyCares practices reported continuing to foster relationships with SNFs and nursing homes to help coordinate care. In addition, to continue to strengthen existing relationships,

[1] Information in this appendix 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. See Appendix E for more information.

[2] 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.

clinicians frequently reached out to these care partners to remind them to coordinate with the practice when caring for their IAH beneficiaries.

Each HarmonyCares 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

Seven IAH practices (Boston, Cleveland, Long Island, Philadelphia, Richmond, Washington, and Wilmington) were part of nonprofit academic medical centers or health systems with academic missions. Three of these organizations comprised the Richmond-based consortium (Philadelphia, Richmond, and Baltimore/Washington).[3] 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 this section, we present qualitative data for each component of the consortium separately.

All academic medical center practices except for Boston employed a mix of physicians, NPs, and/or PAs to conduct house home visits. Social workers were also 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 two academic medical center practices 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 B.5). 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 center practices varied in their use of technologies to facilitate care delivery and planning. All but two practices relied on clinical judgment to determine the level of care rather than using a formal risk-stratification system. All were notified automatically of patients' hospital admissions or ED visits from at least some hospitals with which they built relationships.

[3] Three organizations (Philadelphia, Richmond, and Baltimore/Washington) participated as one consortium, which the demonstration considers as one practice for the purpose of calculating incentive payments.

2.3. Independent practices

The demonstration included four independent practices (Austin, Brooklyn, Durham, and Portland) that were diverse in size, structure, and operating practices. All four independent practices had staff dedicated to coordinating care for patients; however, the type of staff used to coordinate care varied across the practices. For example, some had nurse care managers, whereas others trained medical assistants or similar staff as patient care coordinators.

The sites of care for independent practices varied across them—Austin and Brooklyn conducted most visits in private home settings, whereas Durham and Portland conducted most visits in assisted living or other group living facilities. Some of the independent practices provided nonbillable visits by social workers and nurse care managers.

Independent practices had staff dedicated to coordinating care for patients, but the type of staff used to coordinate care varied across the practices. For example, some used nurse care managers, and others trained medical assistants or similar staff as patient care coordinators. One independent practice reported risk-stratifying patients as a way to determine the intensity of care the practice would provide, one rolled out risk stratification for a portion of patients only, and the remaining two reported relying on clinicians' judgment for these determinations. These practices reported using different methods for learning about patient hospital admissions and ED visits, such as relying on patients and caregivers, assisted living facilities, or other sites to notify the practice, or receiving notice through health information exchanges.[4]

[4] Health information exchanges allow the electronic sharing of health care information. They can be implemented at different levels, including a region (such as the greater Washington, DC, area), community, or hospital system.

Exhibit B.4. Structural characteristics of IAH practices, as of 2019 (IAH Year 6)

Practice	Affiliation	Ownership	Participates in ACO (years)	Accepts Medicare Advantage plans	Clinicians making home visits			Other staff involved in care team				
					Physicians	PAs	NPs	Care coordina-tors[a]	RNs	MAs	SWs	Others
HarmonyCares Medical Group 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 center practices												
Boston, MA	Boston Medical Center	Nonprofit	Yes (5)	Yes	6 PT	-	2 FT	-	5	-	-	
Cleveland, OH	Cleveland Clinic	Nonprofit	Yes (1)	NA	7 FT	1 PT	3 FT	-	3	4	1	1 nurse manager, 1 pharmacist
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
Baltimore, MD/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 B.4 (*continued*)

Practice	Affiliation	Ownership	Participates in ACO (years)	Accepts Medicare Advantage plans	Clinicians making home visits			Other staff involved in care team				
					Physicians	PAs	NPs	Care coordina-tors[a]	RNs	MAs	SWs	Others
Independent practices												
Austin, TX	Kindred Health Care	For profit	No	NA	4 FT 2 PT	4 FT	9 FT	2	-	2	-	5 LPNs
Brooklyn, NY	Heal[c]	For profit	Yes (5)	Yes	3 FT	7 FT	21 FT	-	-	-	-	1 quality assurance nurse
Durham, NC	None	For profit	No	Yes	35 FT	55 FT[d]	-	-	-	-	12	6 podia-trists, 2 psycholo-gists
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 Year 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 Year 6 of the IAH demonstration. For information on the three practices that left the demonstration after Year 7 (Brooklyn, Portland, and Wilmington), refer to the evaluation report covering Year 7 of the IAH demonstration.

[a] Care coordinators are health professionals who help manage a patient's care by monitoring and coordinating patients' care plans, connecting them with clinicians, and making telephone check-in calls. IAH practices 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 practice did not provide a breakdown of physicians, PAs, or NPs.

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

[d] The Portland practice 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; NA = not available; NP = nurse practitioner; PA = physician assistant; PT = part time; RN = registered nurse; SW = social worker.

Exhibit B.5. Operational characteristics of IAH practices, as of 2019 (IAH Year 6)

Practice	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
HarmonyCares Medical Group practices								
Dallas, TX	9 or 10	175	Yes: nurse navigator	Occasionally[d]	No	88.3	11.1	No
Flint, MI	8 or 9	175	Yes: nurse navigator	Yes	No	77.6	22.4	No
Jacksonville, FL	8 or 9	175	Yes: nurse navigator	Occasionally[d]	No	68.0	32.0	No
Lansing/Ann Arbor, MI	8 or 9	175	Yes: nurse navigator and home health company	Occasionally[d]	No	65.5	34.3	No
Milwaukee, WI	8 or 9	175	Yes: nurse navigator	Yes	No	27.9	70.6	No
Academic medical center practices								
Boston, MA	4	90	Yes: nurse care manager visits	No	Yes: for urgent visits only; uncommon	79.5	4.4	Yes
Cleveland, OH	6 or 7	150 to 200	No	Yes	Yes: for urgent visits only; uncommon	82.7	15.2	No
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 or 4	40	Yes: nurse	No	No	96.1	0.4	Yes
Baltimore, MD/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

Exhibit B.5 (*continued*)

Practice	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
Independent practices								
Austin, TX	10	200	No	No	Yes: for urgent visits only; uncommon	86.9	13.1	No
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
Durham, NC	10 to 15	150 to 200	No	Yes	No	9.4	86.9	No
Portland, OR	4	122	Yes: RN, social worker, or chaplain	Yes	No	16.0	83.9	No

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 Year 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 Year 6 of the IAH demonstration. For information on the three practices that left the demonstration after Year 7 (Brooklyn, Portland, and Wilmington), refer to the evaluation report covering Year 7 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% 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 practice defines it as "Saturdays occasionally." The Jacksonville practice defines it as "a case-by-case basis up to 6/7 times a year." The Lansing/Ann Arbor practice defines it as "up to the providers' discretion."

ALF = assisted living facility; RN = registered nurse; SW = social worker.

Exhibit B.6. Care delivery processes among IAH practices, as of 2019 (IAH Year 6)

Practice	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
HarmonyCares Medical Group 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.
Academic medical center practices				
Boston, MA	No. Clinical judgment only.	Yes	Automated notices from some sites through EHR	Yes. Call as needed based on care plan.
Cleveland, OH	No. Clinical judgment only.	No	Automated notices from some sites through EHR	Yes. Call twice weekly.

Exhibit B.6 (*continued*)

Practice	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
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
Baltimore, MD/ Washington, DC	No. Clinical judgment only.	Yes	Automated notices from some sites through EHR	Yes. Call monthly.
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				
Austin, TX	Yes. Assessment scores, hospital admission history, and clinical judgment used to assign level-of-risk score, which determines level of proactive outreach and care team involvement.	Yes	Rely on patient or caregivers to notify practice	Yes. Call weekly or biweekly based on acuity of patient
Brooklyn, NY	No. Clinical judgment only.	Yes	Automated notice from some sites	Yes. Proactive outreach regarding vaccines in partnership with Merck.
Durham, NC	No. Clinical judgment only.	Yes	Notification from ALFs where majority of patients live	Yes. No formal schedule, call as needed.
Portland, OR	Yes. Not fully rolled out, but it covered about three-quarters of patients in late 2019.	Yes	Automated notice from HIE	Yes. Call as needed based on acuity of patient and if patient was recently hospitalized.

Exhibit B.6 (*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 Year 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 Year 6 of the IAH demonstration. For information on the three practices that left the demonstration after Year 7 (Brooklyn, Portland, and Wilmington), refer to the evaluation report covering Year 7 of the IAH demonstration.

ALF = assisted living facility; ED = emergency department; EHR = electronic health record; HIE = health information exchange.

3. Description of IAH beneficiaries

Beneficiaries who received HBPC from the IAH practices were eligible for the demonstration if they met several criteria related to their health and use of health care (Exhibit B.7). Congress limited the demonstration to 10,000 beneficiaries in each of the first five years, 15,000 beneficiaries in Years 6 and 7, and 20,000 beneficiaries in Years 8 to 10.

Exhibit B.7. Requirements for beneficiaries to be eligible for the IAH demonstration

- Have at least two chronic conditions
- Require human assistance with at least two activities of daily living
- Have been hospitalized and received acute or subacute rehabilitation services in the prior 12 months
- Be enrolled in fee-for-service Medicare
- Not be in long-term care or hospice at the time of enrollment in the demonstration

4. Hypotheses about how IAH may have affected outcomes

Exhibit B.8 provides proposed hypotheses for how HBPC from IAH practices, along with the IAH demonstration payment incentive, may have affected Medicare spending, hospital use, and health outcomes.

Exhibit B.8. Hypotheses of how HBPC from IAH practices, along with the IAH demonstration payment incentive, may have affected outcomes

Features	Hypotheses	Mechanisms
IAH demonstration component: IAH practices' general approach to HBPC		
Provides access to primary care at home for beneficiaries who have limited mobility or costly or unreliable transportation	IAH clinicians identify and address acute issues as they arise and help beneficiaries manage their chronic conditions effectively at home, increasing primary care visits and possibly decreasing specialist visits, which may lead to more spending among clinicians and suppliers. If primary care clinicians take a larger role in beneficiaries' health care, that may reduce spending on imaging and lab services (by reducing duplication of imaging and lab services ordered by other clinicians) or increase spending on imaging and lab services (by identifying previously unmet needs). Identifying and addressing acute and chronic conditions through improved access to primary care may also reduce hospital use, including total and potentially avoidable hospital admissions and ED visits, along with inpatient, outpatient, and total spending. Preventing hospital admissions through increased access to primary care at home also reduces the risk of deconditioning (a decline in physical function) and other negative health outcomes while in the hospital. Avoiding deconditioning in the hospital may prevent permanent loss of function and reduce use of institutional long-term care. Care teams conduct regular home visits to monitor patients closely, which may prevent or shorten SNF stays after hospital use, thereby reducing SNF spending, while possibly increasing spending on home health services. Fewer or shorter hospital stays and SNF stays would lead to more days at home for beneficiaries, thereby improving patient experience.	Traveling to a physician's office presents substantial physical and financial burden for some beneficiaries and their caregivers, which may lead to some beneficiaries having fewer primary care visits than they would otherwise have. IAH practices eliminate this burden by visiting beneficiaries in their home, as well as arranging other in-home services, such as x-rays, ultrasound exams, and blood draws. Beginning HBPC may increase IAH beneficiaries' use of primary care and related services, which leads to timely identification and management of acute issues and exacerbations of chronic conditions.

Exhibit B.8 (*continued*)

Features	Hypotheses	Mechanisms
Allows the clinician to obtain information not obtained in an office visit that may improve health care, avoid accidents, or address HRSN	IAH clinicians identify and address beneficiaries' challenges in managing medications and avoiding accidents, which may reduce hospital use and inpatient, outpatient, and total spending. Having a deeper understanding of a beneficiary's needs could lead to more primary care visits and more spending on home health use and durable medical equipment to meet these needs. Clinicians identify beneficiaries' caregiver support and HRSN, and care teams provide referrals to community-based support services to help meet beneficiaries' needs and improve adherence to care plans. Together, these may reduce hospital use along with inpatient, outpatient, and total spending. Also, they may reduce SNF spending and lead to more days at home for beneficiaries.	During home visits, clinicians learn how beneficiaries store and manage their medications, understand obstacles beneficiaries face in adhering to care plans to manage their conditions, observe caregiver support and interactions with beneficiaries, and identify safety improvements that reduce the risk of falls. Observing beneficiaries in their home also gives clinicians valuable information to help beneficiaries with HRSN, including by providing referrals to needed support services and home and community-based services via Medicaid.
Encourages development of a trusting relationship and effective communication among the beneficiary, caregiver, and clinician	IAH clinicians and beneficiaries communicate effectively and work together to address needs and emerging issues. This may lead to beneficiaries and caregivers contacting the IAH practice before going to the ED if they are unsure whether their condition is an emergency, which may reduce ED visits and outpatient spending. It may also help beneficiaries better manage their chronic conditions, which may reduce hospital admissions, inpatient spending, and total spending. It may lead to clinicians better understanding beneficiaries' goals of care, which may increase days at home and hospice use and reduce use of institutional long-term care if those are consistent with beneficiaries' preferences.	Over time, interacting with beneficiaries and their caregivers in the home helps clinicians build trust with beneficiaries and understand their needs and preferences for care. A trusting relationship between beneficiaries and clinicians may encourage beneficiaries to share information about emerging issues with their care team. In addition, clinicians become aware of and respond promptly to acute exacerbations of chronic conditions.

Exhibit B.8 (*continued*)

Features	Hypotheses	Mechanisms
Coordinates care with home health agencies for beneficiaries receiving home health services	IAH practices identify beneficiaries' needs for home health services and coordinate with home health agencies to meet beneficiaries' needs for additional support for acute problems and chronic condition management. Appropriate use of home health services may increase spending on home health but may also reduce total spending by preventing hospital and ED use. Home health staff monitor beneficiaries at home and coordinate with IAH practices to manage acute issues promptly and effectively, which may allow beneficiaries to shorten or avoid costly SNF stays after hospital use, thereby reducing SNF spending. Fewer or shorter hospital stays and SNF stays would lead to more days at home. Frequent visits from a home health agency offer additional in-home care to complement primary care visits and may reduce loneliness and feelings of social disconnection, which are associated with an increased risk of sickness and death. Thus, IAH may reduce the probability of death in a given period of time.	Home health services refer to services provided under the Medicare home health benefit, which requires a beneficiary to be homebound and needing at least one of the following: (1) intermittent skilled nursing care or (2) physical therapy, speech-language pathology, or occupational therapy services. IAH practices reported having close relationships with home health agencies, and communication and coordination with home health agencies were important parts of preventing or responding to acute problems. Beginning HBPC may increase IAH beneficiaries' spending on home health services.
IAH demonstration component: opportunity to earn performance-based incentive payments for providing HBPC annually		
Tracks beneficiaries across settings, as required by the IAH demonstration, to provide follow-up contact within 48 hours of hospital and ED use	IAH practices receive timely notification of beneficiary hospital and ED visits and provide timely post-discharge follow-up, which may increase primary care visits and help prevent unplanned readmissions. If beneficiaries, caregivers, and other clinicians are aware of these transitional care services, their awareness could also lead to less SNF spending and more days at home.	Early in the IAH demonstration, many IAH practices added staff, such as nurse case managers, to their care teams to track beneficiaries across settings. Some practices expanded their use of electronic health records or electronic health information exchanges to improve timely notification and follow-up of hospital and ED visits. Tracking beneficiaries across settings may allow IAH practices to better address post-discharge needs such as medication changes and a decline in physical function (deconditioning) due to lack of movement while in the hospital.

Exhibit B.8 (*continued*)

Features	Hypotheses	Mechanisms
Offers access to the primary care team at all hours, as required by the IAH demonstration	IAH practices provide 24/7 access to care teams, which allows beneficiaries to obtain help in managing acute issues and avoid ED visits, perhaps by increasing primary care visits and clinician and supplier spending.	Early in the IAH demonstration, IAH practices reported efforts to improve consistency of access at all hours and coordinate after-hours care through communication supported by an electronic health record. Beneficiaries and caregivers can contact care teams 24/7 to help address urgent care needs at home.

ED = emergency department; HRSN = health-related social needs; SNF = skilled nursing facility.

Appendix C.

Technical Appendix for the Performance-Year Analysis

1. Overview of the impact analysis

The IAH demonstration tested whether offering HBPC practices the opportunity to earn performance-based incentive payments for providing HBPC reduced spending and hospital use and improved health outcomes. Thus, the demonstration comprised two distinct interventions: IAH practices' general approach to providing HBPC and the demonstration's payment incentive.

As we describe in Chapter 1 of the report, we conducted two main analyses to understand whether IAH met its goals of reducing spending, reducing hospital use, and improving health outcomes. The evaluation was required to produce estimates of the effect of the demonstration for each performance year to align with the timing of annual incentive payments. We refer to that analysis as the performance-year analysis, and is the focus of this Appendix C. It examines the effects of the minor improvements to care delivery for beneficiaries receiving HBPC made by IAH practices because of the IAH payment incentive. As with the annual incentive payment calculation, the performance-year analysis included all IAH-eligible beneficiaries of IAH practices—those who were new beneficiaries and those who were existing HBPC beneficiaries of the IAH practice at the time they met IAH eligibility criteria. Because IAH practices had been providing HBPC to some beneficiaries before the demonstration began in 2012, the only way to examine the effects of IAH practices' general approach to providing HBPC was to limit the sample to new beneficiaries of IAH practices. We refer to that study as the three-year analysis of IAH beneficiaries new to HBPC, and we describe it in Appendix D.

Exhibit C.1 describes what the difference-in-differences estimate from the performance-year analysis measures. The change from before the demonstration to during the demonstration that differed for IAH and comparison beneficiaries represents the effect of the opportunity for HBPC practices to earn IAH incentive payments. The effect of this payment incentive is what we estimate with the difference-in-differences analysis we describe in Chapter 5 of this appendix.

Exhibit C.1. Results for outcomes estimated by the performance-year analysis

Intervention received before the demonstration (pre-period)			Intervention received after the demonstration began (post-period)			Intervention tested by the difference-in-differences analysis
IAH beneficiaries (1)	Comparison beneficiaries (2)	IAH-Comparison difference (1) – (2) (3)	IAH beneficiaries (4)	Comparison beneficiaries (5)	IAH-Comparison difference (4) – (5) (6)	(6) – (3) (7)
Received HBPC from an IAH practice during a time that the practice could *not* have earned IAH incentive payments	Did not receive HBPC, and the beneficiary's provider could *not* have earned IAH incentive payments	Receipt of HBPC from IAH practice	Received HBPC from an IAH practice during a time that the practice could have earned IAH incentive payments	Did not receive HBPC, and the beneficiary's provider could *not* have earned IAH incentive payments	Receipt of HBPC from IAH practice and opportunity to earn IAH incentive payments	The DID estimate reflects the effect of the opportunity for HBPC practices to earn IAH incentive payments

2. Methods for identifying the IAH beneficiaries

To comply with the legislation that established the IAH demonstration (see Appendix A), the demonstration used a practice-based enrollment process. Practices 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 practices 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 practices 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 4 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 practices 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 practices 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 each year. After explaining these approaches in Sections 2.1 and 2.2, we provide details about reasons for differences in the counts of IAH practices' eligible beneficiaries in Section 2.3.

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

The IAH practices identified beneficiaries they thought were eligible to participate in the demonstration; we list the eligibility requirements in Appendix B. After providing these beneficiaries with information about the demonstration and visiting homes to explain it, the IAH practices 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 practices identify potential beneficiaries for enrollment into the demonstration based on the eligibility criteria. Each year, 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 practice 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 the Centers for Medicare & Medicaid Services (CMS) during the Public Health Emergency. The contractor provided the practices with information on the potential enrollees, and the practices then reviewed their records and assessed additional information about the beneficiaries' eligibility (such as whether they met the activity of daily living [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 or telephone visit by the practice may 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 practices provided the implementation contractor with the reason for the beneficiary's ineligibility. Reasons practices 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 their 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 fee-for-service (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.

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

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 practices. 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, as measured by chronic conditions indicators in the Chronic Conditions Warehouse
- Inpatient hospital admission or observation stay in the previous 12 months[5]
- Acute or subacute rehabilitation services use in the previous 12 months[6]
- 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, Year 9 was January to December 2022. If a beneficiary had a hospital admission in July 2022 and home health care in October 2022, that person would be eligible for demonstration Year 9 as of November 1, 2022.

In the following section, we explain why we removed from the sample IAH practices that did not participate in the demonstration in a given year. Next, we 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.[7]

2.2.1. 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 and our estimates for Year 6 are based on data for the 12 practices that completed Year 6. Similarly, our estimates for Years 7 to 10 are based on data for the practices that completed Years 7, 8, 9, and 10 (10 practices in Year 7, seven practices in Years 8 and 9, and one practice in Year 10). 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

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

[6] 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.

[7] For more information about reasons for the differences between demonstration enrollment and evaluation, refer to the evaluation report covering Year 4 of the IAH demonstration.

demonstration in Year 2 did not have enough home-based primary care beneficiaries 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 across the entire demonstration— stopped providing home-based primary care after it left the demonstration after Year 5.[8] 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 5 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 HBPC to Medicare beneficiaries with multiple chronic conditions and substantial functional limitations. Rather, among the pool of HBPC practices that volunteered for the demonstration, CMS selected 18 practices 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 HBPC from practices other than those in the demonstration. See Chapter 1 of the report for more information about how these evaluation results cannot be generalized to other HBPC providers or other beneficiaries.

2.2.2. 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 assistance from another person with the activity. Assistance from a device such as a cane for walking or grab bar for bathing did not qualify as an ADL limitation for enrollment in the IAH demonstration. The exception to this general guideline was for wheelchair use: use of a wheelchair as the primary mode of mobility with or without assistance from another person qualified as an ADL limitation.

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 from home health agencies (not home-based primary care visits from primary care clinicians); (2) the Minimum Data Set (MDS), collected when beneficiaries receive skilled nursing facility (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 toilet. Each assessment instrument has one or more data elements that indicate

[8] For more information, refer to the evaluation report covering Year 5 covering Year 5 of the IAH demonstration.

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 beneficiaries. 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 since we could not determine which beneficiaries had human assistance and who needed just device assistance. 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.

However, 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.[9]

[9] 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.

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.

2.2.3. Attribution and enrollment data

In addition to determining eligibility for the demonstration, in each year we applied the following criteria for attributing a beneficiary to a demonstration practice (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. Beginning 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, they had at least one additional visit from the demonstration practice in the home, in an assisted living facility, in an office, or by telephone.

The demonstration rules required that all beneficiaries 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 practice as Mathematica-eligible IAH beneficiaries (or simply IAH beneficiaries). IAH beneficiaries were the intervention group for the evaluation. For beneficiaries to be in the IAH group for the evaluation, they

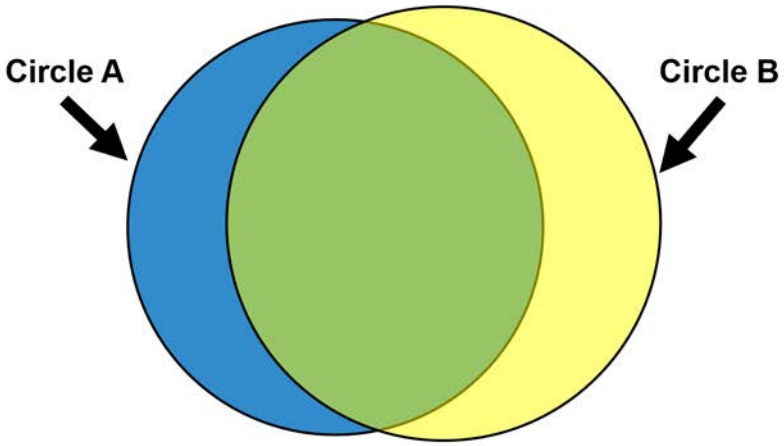
had to meet the eligibility and attribution criteria outlined above according to Mathematica's analysis of Medicare enrollment, claims, and assessment data.

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 practice according to administrative data (in Exhibit C.2, 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 C.2 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 intervention 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 practice according to administrative data.

On average across the 10 demonstration years, about one-quarter of IAH beneficiaries each year were new patients of the IAH practice and new to HBPC more generally. These beneficiaries had not received care from an IAH practice or HBPC from any clinician in the two years before becoming eligible for the demonstration. The percentage of IAH beneficiaries who were new to HBPC in each year varied a bit, ranging from 24% to 26% in the year before the demonstration and Years 1 to 6 to 19% to 20% in Years 7 to 9 (during the COVID-19 pandemic). It was 23% in Year 10, when one practice participated in the demonstration.

Exhibit C.2. 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 2022 (Month 2 of Year 9) and moved out of the IAH practice's geographic area or entered long-term care in September 2022, we continued to follow that beneficiary through the end of the study year (December 31, 2022).

Demonstration Year 1 (June 2012 to May 2013).[10] Mathematica identified 8,216 beneficiaries who met the demonstration eligibility criteria and were attributed to the demonstration practice during the first year (Exhibit C.3). 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 C.2, 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 C.2, 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 whom we did not find eligible for the demonstration using administrative data (in Exhibit C.2, Circle A excluding Circle B—the blue crescent).

[10] For practices that began the demonstration in June 2012, Month 1 was June. For practices that began the demonstration in September 2012, Month 1 was September.

Exhibit C.3. Numbers of beneficiaries based on different identification processes

Demonstration year	Mathematica-eligible IAH beneficiaries[a]			IAH-enrolled only (blue crescent)	Number of practices in the evaluation
	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	14
2	4,564	2,702	7,266	4,059	14
3	4,498	3,066	7,564	4,718	14
4	6,019	3,485	9,504	5,663	14
5	5,950	4,008	9,958	6,407	14
6	2,692[b]	5,301[b]	7,993	3,696[b]	12
7	2,335	2,723	5,058	4,101	10
8	1,547	1,795	3,342	2,354	7
9	819	1,801	2,620	1,917	7
10	125	42	167	283	1

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

[a] This corresponds to the yellow circle in Exhibit C.2, 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 beneficiaries 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).[11] 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 C.2, 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 C.2, the blue crescent) included people who enrolled for the 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

[11] For practices that began the demonstration in June 2012, Month 1 was June. For practices that began the demonstration in September 2012, Month 1 was September.

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 C.2, 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).[12] 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 beneficiaries who met the eligibility criteria or beneficiaries who previously received care from the IAH practice and did not meet the eligibility criteria for the demonstration until Year 3.

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.

The IAH group was substantially larger in Year 4 than in previous demonstration years. For all practices combined, this group increased 26 percent from Year 3 to Year 4. Five practices had increases of more than 20 percent from Year 3 to Year 4: Brooklyn, Durham, Dallas, Flint, and Portland. This increase could have reflected the expansion of existing IAH practices. Brooklyn merged with another HBPC practice, and the Durham practice expanded throughout North Carolina during the demonstration. In Year 4, Dallas expanded into a new geographic area, and Flint added clinicians in its existing geographic area.

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

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 beneficiaries to the IAH practice. We explore the implications of ACO participation in Chapter 6 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.

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 HBPC 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.

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 after Year 5, (2) a substantial decrease in the number of IAH beneficiaries identified for Brooklyn, and (3) a drop in the number of clinicians treating beneficiaries 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 finalized 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 practice-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—posed a substantial risk of bias to the estimated impacts of the demonstration in Year 6. This is because these issues affected 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.

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 practice; other than Brooklyn, the number of IAH beneficiaries at each practice 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 practices 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 for the group of 10 practices that participated in Year 7, 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 among the same group of practices in Year 5.

Second, we included IAH-eligible beneficiaries cared for by 12 clinicians (nine in Brooklyn and three 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 practices 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 practices 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 time frame in Year 6. This decrease was due to (1) less SNF and nursing facility use during the COVID-19 pandemic in 2020 than in previous years and (2) CMS temporarily waiving MDS assessment time frame requirements during the COVID-19 public health emergency for SNFs and allowing states to waive the time frame requirements for nursing facilities if needed.[13] Concerns about the risk of contracting COVID-19 may 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 time frame requirement, it noted that most facilities completed assessments in a timely manner.[14] 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 time frame waiver.

To assess how the change in ADL data may 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 C.4). 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

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

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

percentage points for comparison beneficiaries. The missing assessment data probably did not pose a substantial risk of bias to the estimated impacts of the demonstration in Year 7. These issues affected both groups, 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 C.4. Percentage of beneficiaries identified as needing assistance with two or more ADLs, by assessment data source in Years 6 and 7 among practices participating in Year 7

	IAH			Comparison[a]		
Assessment	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 3 of this appendix for details regarding the matched comparison beneficiary selection process.

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

ADLs = 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 beneficiaries of IAH practices were more likely to use home health at all and used home health for more days than comparison beneficiaries.[15] 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 with expenditures and other outcomes during the COVID-19 pandemic (Year 7) than in earlier years.

[15] 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 Year 4 of the IAH demonstration.

Demonstration Year 8 (January to December 2021). In Year 8, Mathematica identified 3,342 beneficiaries who met the demonstration eligibility criteria and were attributed to the demonstration practices. Of these 3,342 IAH beneficiaries, 1,547 were enrolled in the demonstration, and 1,795 were not. These 3,342 beneficiaries included people in the analysis sample in Years 1 to 7 who requalified in Year 8 by meeting eligibility and attribution requirements, as well as those not in the analysis sample in any of the first seven years.

The number of beneficiaries we attributed to the IAH group using administrative data decreased by 34 percent from Year 7 to Year 8 (Exhibit C.5). Much of this decrease can be attributed to Brooklyn, Portland, and Wilmington leaving the demonstration after Year 7. Excluding those three practices, the decrease from Year 7 to Year 8 was 20 percent.

One factor that contributed to the decrease in the number of IAH beneficiaries is decreasing enrollment in Medicare FFS (because of increasing enrollment in Medicare Advantage). A second factor is the large decrease in the number of clinicians associated with the practices. There was a 24 percent decrease in the number of clinicians from Year 7 to Year 8, leaving only 109 active clinicians whose identification numbers we used to identify IAH beneficiaries in Year 8.

Exhibit C.5. IAH beneficiaries and clinicians in Years 6, 7, and 8, by practice

Practice	IAH beneficiaries					IAH clinicians				
	Year 6	Year 7	Percentage change, Year 6 to Year 7	Year 8	Percentage change, Year 7 to Year 8	Year 6	Year 7	Percentage change, Year 6 to Year 7	Year 8	Percentage change, Year 7 to Year 8
Long Island, NY	331	262	-21	222	-15	16	16	0	14	-13
Dallas, TX	1,290	1,121	-13	792	-29	23	22	-4	19	-14
Flint, MI	1,415	1,181	-17	978	-17	38	29	-24	22	-24
Jacksonville, FL	621	499	-20	402	-19	13	13	0	8	-38
Lansing, MI	608	458	-25	407	-11	26	15	-42	11	-27
Milwaukee, WI	450	394	-12	352	-11	12	13	8	10	-23
Richmond, VA (3 organizations)[a]	310	253	-18	189	-25	36	36	0	25	-31
Total	**5,025**	**4,168**	**-17**	**3,342**	**-20**	**164**	**144**	-12	**109**	**-24**

Source: Mathematica's analysis of data from the IAH implementation contractor and 2018–2021 Medicare claims and enrollment data from the Chronic Conditions Warehouse for IAH practices that participated in Year 8.

[a] The Washington, DC-based component of the Baltimore/Washington organization that was a member of the consortium left the demonstration after Year 7 ended.

A final factor that may have contributed to identifying fewer IAH beneficiaries is COVID-19 pandemic-related changes in the availability of the ADL assessment data we used to identify beneficiaries eligible for the demonstration. When we examined assessment data for all states (not just states where IAH practices were located), we identified about the same number of beneficiaries who needed help from another person with two or more ADLs in Year 8 as in Year 7 (an increase of 1.3 percent). This followed a 10 percent decrease from Year 6 to Year 7. The similar number of beneficiaries identified with two or more ADLs in Years 7 and 8 suggests that factors that caused the decrease in the number of beneficiaries who met the ADL requirement—changes in health care use during the pandemic and changes CMS made in response to the public health emergency—continued to affect use of rehabilitation services in Year 8.

As we described earlier in this chapter, we use three ADL assessment data sources to identify IAH-eligible beneficiaries: OASIS, MDS, and IRF-PAI. To explore whether changes in ADL data from pre-pandemic years may have affected sample identification, we examined the percentage of beneficiaries we identified as needing assistance from another person with two or more ADLs by assessment type (Exhibit C.6). The percentage identified by MDS only increased slightly since the first year of the COVID-19 pandemic for both the IAH and comparison groups, while the percentage identified by OASIS only decreased slightly. More importantly, these data show that the share of beneficiaries identified by OASIS only was more than 40 percent higher for the IAH group than the comparison group (57 percent of IAH beneficiaries and 40 percent of comparison beneficiaries in Year 8). At the same time, the share identified by the MDS only was two-thirds lower for the IAH group (12 percent of IAH beneficiaries and 33 percent of comparison beneficiaries). This was also true in Year 6, before the pandemic began (Exhibit C.4). As we discussed earlier in this section, these differences suggest that changes in unobservable differences between IAH and comparison beneficiaries could pose a risk of bias during the pandemic.

Exhibit C.6. Percentage of beneficiaries identified as needing human assistance with two or more ADLs, by assessment data source in Years 7 and 8 among practices participating in Year 8

	IAH			Comparison[a]		
Assessment	Year 7[b]	Year 8	Percentage point difference	Year 7[b]	Year 8	Percentage point difference
OASIS only	59.6	57.0	-2.6	42.6	39.5	-3.1
MDS only	10.4	11.8	1.5	31.3	33.2	1.9
IRF-PAI only	0.7	0.9	0.2	2.8	3.2	0.4
MDS and IRF-PAI	0.5	0.6	0.1	1.6	1.7	0.1
OASIS and MDS	24.0	24.6	0.6	17.0	17.5	0.5
OASIS and IRF-PAI	2.6	2.9	0.3	3.0	3.1	0.1
OASIS, MDS, and IRF-PAI	2.2	2.1	-0.1	1.8	1.8	0.0

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 3 of this appendix for details regarding the matched comparison beneficiary selection process.

Exhibit C.6 (*continued*)

[b] Year 7 data are restricted to practices that also participated in Year 8.

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

Demonstration Year 9 (January to December 2022). In Year 9, Mathematica identified 2,620 beneficiaries who met the demonstration eligibility criteria and were attributed to the demonstration practices. Of these 2,620 IAH beneficiaries, 819 were enrolled in the demonstration, and 1,801 were not. These 2,620 beneficiaries included people in the analysis sample in Years 1 to 8 who requalified in Year 9 by meeting eligibility and attribution requirements, as well as those not in the analysis sample in any of the first eight years.

The number of beneficiaries we attributed to the IAH group using administrative data decreased by 22 percent from Year 8 to Year 9 (Exhibit C.7). The number of IAH beneficiaries at each practice decreased by between 12 and 40 percent. This followed decreases from Year 7 to Year 8 that ranged from 11 percent to 29 percent.

One factor that contributed to the decrease in the number of IAH beneficiaries is decreasing enrollment in Medicare FFS (because of increasing enrollment in Medicare Advantage). A second factor is a decrease in the number of clinicians associated with the practices. There was a 13 percent decrease in the number of clinicians from Year 8 to Year 9, leaving only 95 active clinicians whose identification numbers we used to identify IAH beneficiaries in Year 9. A similar decrease occurred from Year 7 to Year 8.

Exhibit C.7. IAH beneficiaries and clinicians in Years 8 and 9, by practice

Practice	IAH beneficiaries			IAH clinicians		
	Year 8	Year 9	Percentage change, Year 8 to Year 9	Year 8	Year 9	Percentage change, Year 8 to Year 9
Long Island, NY	222	133	-40%	14	14	0%
Dallas, TX	792	646	-18%	19	17	-11%
Flint, MI	978	754	-23%	22	17	-23%
Jacksonville, FL	402	353	-12%	8	6	-25%
Lansing, MI	407	320	-21%	11	8	-27%
Milwaukee, WI	352	248	-30%	10	8	-20%
Richmond, VA (3 organizations)[a]	189	166	-12%	25	25	0%
Total	**3,342**	**2,620**	**-22%**	**109**	**95**	**-13%**

Source: Mathematica's analysis of data from the IAH implementation contractor and 2020–2022 Medicare claims and enrollment data from the Chronic Conditions Warehouse for IAH practices that participated in Years 8 and 9.

[a] The Washington, DC-based component of the Baltimore/Washington organization that was a member of the consortium left the demonstration after Year 7 ended.

A final factor that may have contributed to identifying fewer IAH beneficiaries in Year 9 (as in Years 7 and 8) is COVID-19 pandemic-related changes in the availability of the ADL assessment data we used to identify beneficiaries eligible for the demonstration. When we examined assessment data for all states (not just states where IAH practices were located), we found that the number of beneficiaries identified with two or more ADLs in all years affected by the COVID-19 pandemic were considerably lower than in Year 6. The number of beneficiaries who needed help from another person with two or more ADLs in Year 9 (3.57 million) was slightly higher than in Years 7 and 8 (3.24 million and 3.39 million, respectively) but still much lower than in Year 6 (4.10 million). This suggests that changes in health care use during the pandemic and changes CMS made in response to the public health emergency in Year 7 continued to affect use of rehabilitation services in Years 8 and 9—particularly SNFs and nursing homes where MDS data are gathered. As we discussed earlier in this section, these differences suggest that changes in unobservable differences between IAH and comparison beneficiaries could pose a risk of bias during the pandemic.

Demonstration Year 10 (January to December 2023). In Year 10, Mathematica identified 167 beneficiaries who met the demonstration eligibility criteria and were attributed to the one practice that participated in Year 10. Of these 167 IAH beneficiaries, 125 were enrolled in the demonstration, and 42 were not.

2.3. Implications of the statutory limit on enrollment

The maximum number of beneficiaries in the IAH demonstration was 10,000 in the legislation that established IAH. This limit applied to Years 1 to 5. Congress raised the maximum number of beneficiaries to 15,000 starting in Year 6 and 20,000 starting in Year 8. This limit did not apply to the evaluation; however, the number of beneficiaries Mathematica identified as eligible for the demonstration and attributed to an IAH practice never exceeded the limit.

During years that the number of enrollees identified by the implementation contractor exceeded the statutory limit, the implementation contractor modified payments to ensure that the demonstration operated in compliance with the legislation.

3. Methods for 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 HBPC. 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.

3.1. 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 (two pre-demonstration years and 10 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 practices 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 in Year 2. We refer to these comparison groups as potential because we identified the final comparison groups using propensity score matching (described later in this chapter). In some of the markets where IAH practices operated, the number of Medicare FFS beneficiaries who received HBPC from a practice other than an IAH practice would have been quite small, especially after limiting to those who met IAH eligibility criteria. Therefore, we limited the potential comparison group to beneficiaries who did not receive HBPC. We excluded 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 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. Although the comparison group for the performance-year analysis was limited to beneficiaries who did not receive HBPC, the estimates produced by this analysis do *not* reflect the effect of IAH practices' general approach to HBPC on spending and other outcomes. See Chapter 1 of this appendix and Chapter 1 of Appendix D for more information about the difference in the effects estimated by the performance-year analysis and the three-year analysis.

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 practice 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 practice 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 practices located in states that had two demonstration practices (Flint and Lansing in Michigan in Years 1 to 9; Brooklyn and Long Island in New York in Years 1 to 7; and 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.[16] For each pair of practices located in the same state, we included in the preliminary model all of the IAH beneficiaries in those two practices as well as all beneficiaries in the comparison pool for both practices after applying the ZIP code restriction. Each comparison beneficiary was matched to an IAH beneficiary in one of the two practices; this process determined the practice 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 practices.

As with IAH beneficiaries, we again identified beneficiaries who were in the matched comparison group in earlier demonstration years as potential comparison beneficiaries in a later year if they met all IAH eligibility requirements in that later year.

3.2. 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 HBPC. 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 intervention group on key observable characteristics that affect treatment status (such as receipt of HBPC 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 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 9, for example, we matched 2,620 IAH beneficiaries on observable characteristics with beneficiaries who were similar and lived in the same geographic area but did not receive HBPC. We matched each practice separately, including each member of the Mid-Atlantic Consortium (three practices, with the Consortium based in Richmond). We created a comparison group for each practice by estimating a propensity score equation using data for

[16] Conducting a preliminary match typically provides a better match in both practices 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.

the IAH group and the potential comparison group and then used 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 month, two or three months, 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 (Month 1, Months 2 to 6, or Months 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. We included several demographic and enrollment variables as well as the number of ADLs (two, three or four, or five or six) in the matching model but did not seek exact matches for them. We used an indicator variable to identify beneficiaries with missing information for feeding assistance.

We also used various measures of health status in the matching model. We measured individual hierarchical condition category (HCCs) scores 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.[17,18] 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 intervention 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. Beginning in Year 7, we included individual HCCs for morbid obesity and immune deficiency, because they were identified as risk factors

[17] 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.

[18] We used software version V2119 to calculate HCC scores for beneficiaries in Year 8, and incorporated 2021 updates from version 10 of the *International Classification of Diseases*.

for negative outcomes resulting from COVID-19 infection (Bosworth et al. 2021). We also used the HCC risk score itself as a matching variable. Additional details about how we calculated the HCC score and HCC indicators are available in Chapter 4 of this appendix.

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 use COVID-19 diagnosis as a matching variable 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 5 of this appendix, we estimated sensitivity analyses for pandemic years controlling for COVID-19 diagnosis and COVID-19 hospitalization.

More information about variables used in propensity score matching is available upon request.

3.3. Health status, functional status, and demographic characteristics for IAH and matched comparison beneficiaries

The standardized difference in means is a standard statistic used to assess similarities between the intervention 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 variables used in the matching algorithm and variables that might be important to control for but could not be included, such as individual HCCs aggregated with other HCCs in the matching equation (for example, cirrhosis of the liver and chronic hepatitis), and individual chronic conditions measured by the Chronic Conditions Warehouse.

In this appendix, we use Year 5 as an example to demonstrate balance between the IAH and matched comparison beneficiaries. For details on other years, see previous annual reports from the evaluation of IAH. In Year 5, across all 14 practices together (treating the three Mid-Atlantic Consortium practices as one), the absolute value of the standardized difference in the fifth demonstration year was less than 0.10 on all matching variables and less than 0.10 on all but one nonmatching variable (Exhibit C.8). All 14

practices individually had standardized differences of less than 0.25 on all of the matching variables; for 12 of those practices, the standardized differences were also less than 0.10 on all of the matching variables (data not shown). Furthermore, eight of the practices had standardized differences of less than 0.25 on all of the nonmatching variables. As in Year 5, the IAH and matched comparison groups were very similar in the other nine demonstration years.

Exhibit C.8. Characteristics of potential comparison beneficiaries, matched comparison beneficiaries, and IAH beneficiaries, Year 5

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	57.5	38.9	38.9	0.000
Two or three	15.7	17.3	17.3	0.000
Four or more	26.8	43.8	43.8	0.000
Month of the demonstration year in which beneficiary met eligibility criteria, percentage[b]				
Month 1	42.4	65.7	65.2	-0.009
Months 2–6	29.3	21.1	21.5	0.008
Months 7–12	28.3	13.2	13.3	0.004
Proportion with observation stay and no inpatient admission in previous 12 months	5.6	9.5	8.6	-0.031
Demographic characteristics				
Age, percentage				
Proportion younger than 65	12.9	15.6	16.1	0.014
Proportion ages 65–79	42.5	32.8	32.8	0.001
Proportion 80 or older	44.6	51.6	51.1	-0.011
Original reason for Medicare entitlement, percentage				
Proportion whose original eligibility was due to age	73.6	67	66.8	-0.005
Proportion whose original eligibility was due to disability	24.3	31.7	32	0.008
Proportion whose original eligibility was due to ESRD or ESRD plus disability	2.2	1.3	1.2	-0.010
Number of ADLs requiring human assistance, percentage				
Proportion with two ADLs	12.4	7.3	7.2	-0.005
Proportion with three or four ADLs	30.9	31.4	30.7	-0.014
Proportion with five or six ADLs	56.6	61.3	62.1	0.016
Proportion missing information about feeding ADL	9.1	16.6	17.2	0.019

Exhibit C.8 (*continued*)

Variable	Potential comparison group	Matched comparison group	IAH beneficiaries	Standardized difference
Health status				
HCC				
HCC risk score	3.6	4.0	4.1	0.023
Percentage with HCCs				
HCC8: Metastatic cancer	4.6	1.7	1.6	-0.008
HCC9–10: Lung, lymphoma, and other cancers	5.9	3.5	3.5	0.002
HCC11–12: Colorectal, bladder, breast, prostate, and other cancers	10.7	8.6	8.3	-0.011
HCC18: Diabetes with chronic complications	35.9	38.1	37.4	-0.014
HCC21: Protein-calorie malnutrition	13.9	19.7	20.9	0.031
HCC27: End-stage liver disease	1.8	1.2	1.1	-0.005
HCC28–29: Cirrhosis of liver and chronic hepatitis	2.7	2.5	2.4	-0.005
HCC46: Severe hematological disorders	1.8	1.3	1.3	0.001
HCC48: Coagulation defects and other specified hematological disorders	17.2	13.8	13.5	-0.009
HCC51: Dementia with complications	6.8	16.7	17.7	0.028
HCC52: Dementia without complications	19.2	32.2	30.9	-0.028
HCC54–55: Drug/alcohol psychosis and drug/alcohol dependence	7.2	7.7	7.8	0.004
HCC57–58: Schizophrenia, major depressive, bipolar, and paranoid disorders	18.4	27.9	28.3	0.010
HCC70–71: Quadriplegia, paraplegia	2.7	7.1	8.1	0.041
HCC72: Spinal cord disorders/injuries	2.4	1.6	1.6	-0.005
HCC85: Congestive heart failure	45	51.8	51.3	-0.011
HCC96: Specified heart arrhythmias	38.9	36.6	36	-0.014
HCC103–104: Hemiplegia/hemiparesis, monoplegia, other paralytic syndromes	10.2	14.1	14	-0.002
HCC106: Atherosclerosis of the extremities with ulceration or gangrene	4.6	4.6	4.7	0.004
HCC107–108: Vascular disease with or without complications	42.9	50.4	49.9	-0.010
HCC111: Chronic obstructive pulmonary disease	32.9	37.2	37.1	-0.003
HCC134: Dialysis status	5.8	4	3.8	-0.010
HCC136–138: Chronic kidney disease, stages 3–5	9.1	11.1	11.3	0.008

Exhibit C.8 (*continued*)

Variable	Potential comparison group	Matched comparison group	IAH beneficiaries	Standardized difference
HCC139–140: Chronic kidney disease stages 1–2, unspecified renal failure	4.8	6.6	6.7	0.006
HCC157–159: Pressure ulcer of skin with necrosis or skin loss	8.1	16.6	18.1	0.044
Number of chronic conditions measured by Chronic Conditions Warehouse				
Fewer than six	18.5	11.1	11.5	0.012
Six to nine	48.8	46.3	45.7	-0.012
More than nine	32.7	42.5	42.7	0.004
Other measures of health status				
Anemia	18.2	18.7	18.9	0.005
Depression	43.6	56.3	56.7	0.008
Fluid and electrolyte disorders[c]	38.5	40.9	40.4	-0.010
Diagnosis of chronically critically ill or medically complex[d]	32.0	32.1	32.2	0.002
Complication or comorbidity or major complication or comorbidity during the most recent hospital admission	57.4	57.9	57.9	0.001

Source: Medicare claims and enrollment data for 2009–2017 obtained from the Chronic Conditions Warehouse for IAH and matched comparison group beneficiaries in all IAH practices that participated in demonstration Year 5.

Notes: The final sample sizes in Year 5 were 9,958 IAH beneficiaries and 41,387 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 practices in Year 5, Month 1 was October.

[c] Measured using claims from the most recent inpatient and observation stays in the year before the demonstration eligibility date. Diagnosis codes for these conditions were drawn from Gagne et al. (2011).

[d] Measured using diagnoses from the most recent inpatient stay in the year before the demonstration eligibility date. Diagnoses were drawn from Kandilov et al. (2014).

ADLs = activities of daily living; ESRD = end-stage renal disease; HCC = hierarchical condition category.

3.4. 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 10 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 C.9 through C.13). This

difference arose for two reasons. First, the comparison beneficiaries were more likely to die during the demonstration year than the IAH beneficiaries; for example, the rate of death in Year 9 was 17.7 percent for IAH beneficiaries and 25.7 percent for comparison beneficiaries. Second, 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 six months in Year 9 had half the weight of a beneficiary eligible for the demonstration for 12 months in Year 9. 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 5 of this appendix provides additional details about weights and control variables.

Exhibit C.9. Sample for the performance-year analysis, 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: Mathematica's analysis of data from the IAH implementation contractor and Medicare claims and enrollment data from the Chronic Conditions Data Warehouse for 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 C.10. Sample for the performance-year analysis, 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: Mathematica's analysis of data from the IAH implementation contractor and Medicare claims and enrollment data from the Chronic Conditions Data Warehouse for 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 C.11. Sample for the performance-year analysis, 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: Mathematica's analysis of data from the IAH implementation contractor and Medicare claims and enrollment data from the Chronic Conditions Data Warehouse for 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.

Exhibit C.12. Sample for the performance-year analysis, by years, practices that participated in Years 8 and 9

	Two years before the demonstration	One year before the demonstration	Year 1	Year 2	Year 3	Year 4	Year 5	Year 6	Year 7	Year 8	Year 9
Number of IAH beneficiaries	4,294	4,484	5,269	4,226	4,318	5,194	5,490	5,025	4,168	3,342	2,620
Total number of eligible months for IAH beneficiaries	41,519	43,275	51,123	40,782	41,446	50.222	52,809	48,247	40,276	31,048	24,966
Average number of eligible months per IAH beneficiary	9.7	9.7	9.7	9.7	9.6	9.7	9.6	9.6	9.7	9.3	9.5
Number of comparison beneficiaries	18,597	19,436	20,745	18,756	16,993	19,924	22,348	20,765	17,032	13,525	10,703
Total number of eligible months for comparison beneficiaries	166,677	175,955	185,922	171,451	152,174	177,333	196,799	184,314	148,125	111,949	90,842
Average number of eligible months per comparison beneficiary	9.0	9.1	9.0	9.1	9.0	8.9	8.8	8.9	8.7	8.3	8.5

Source: Mathematica's analysis of data from the IAH implementation contractor and Medicare claims and enrollment data from the Chronic Conditions Data Warehouse for all IAH practices that participated in Years 8 and 9

Note: This table reflects the sample we used to estimate the effects of IAH in Years 8 and 9.

Exhibit C.13. Sample for the performance-year analysis, by years, practice that participated in Year 10

	Two years before the demonstration	One year before the demonstration	Year 1	Year 2	Year 3	Year 4	Year 5	Year 6	Year 7	Year 8	Year 9	Year 10
Number of IAH beneficiaries	239	272	246	220	220	235	288	331	262	222	133	167
Total number of eligible months for IAH beneficiaries	2,331	2,576	2,288	2,064	2,077	2243	2,685	3,074	2,549	1,985	1,354	1,655
Average number of eligible months per IAH beneficiary	9.8	9.5	9.3	9.4	9.4	9.5	9.3	9.3	9.7	8.9	10.2	9.9
Number of comparison beneficiaries	932	1,059	1,093	969	984	1,040	1,256	1,497	1,160	991	584	778
Total number of eligible months for comparison beneficiaries	8,187	9,159	9,287	8,601	8,432	9,244	10,940	12,249	9,810	7,556	5,115	6,887
Average number of eligible months per comparison beneficiary	8.8	8.6	8.5	8.9	8.6	8.9	8.7	8.2	8.5	7.6	8.8	8.9

Source: Mathematica's analysis of data from the IAH implementation contractor and Medicare claims and enrollment data from the Chronic Conditions Data Warehouse for the IAH practice that participated in Year 10.

Note: This table reflects the sample we used to estimate the effects of IAH in Year 10.

4. Medicare data and measures used in the impact analysis

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 C.14). We accessed all data through the Chronic Conditions Warehouse Data Enclave.

Exhibit C.14. 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						✓
Area Deprivation Index	✓					

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.

4.1. 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.

In all demonstration and pre-demonstration years, we calculated HCC scores 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 continued using it in Years 7 to 10 (2020 to 2023) rather than switching to the updated Version 24 to maintain consistency across years. CMS-HCC Model V24 is fundamentally different from V21. For example, V24 creates separate models for six segments of the population depending on the beneficiary's reason for Medicare entitlement and Medicaid status. 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 24 could affect IAH and comparison beneficiaries differently, which would have posed 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.

4.2. Dual eligibility for Medicare and Medicaid

When we did propensity score matching for the full sample in all demonstration and pre-demonstration years, we measured dual eligibility for Medicare and Medicaid using the monthly Part A and Part B state buy-in variables on Medicare enrollment data. We used this approach for two reasons. First, Medicaid enrollment data were not available promptly enough for us to define dual eligibility using those data. Second, the evaluation began several years before the Master 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 of a given year (such as IAH Year 1), 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.

4.3. Residence in a more or less disadvantaged area

We used the Area Deprivation Index (ADI) to distinguish between more or less disadvantaged areas (see Chapter 9 of this appendix for results of this analysis). The ADI score ranks neighborhoods by socioeconomic disadvantage at the national level, with a higher score corresponding to a higher level of disadvantage. It includes factors for the theoretical domains of income, education, employment, and housing quality. We define residing in a more disadvantaged area as the beneficiary living in a nine-digit ZIP code with an ADI score at or above the 75th percentile across the United States. Residing in a less disadvantaged area reflects a score of less than 75. In this section, we describe how we constructed this measure in Year 8; the same approach applies to other years of the demonstration and the two pre-demonstration years. In Year 8, 36.4 percent of IAH beneficiaries lived in a more disadvantaged area compared with 32.6 percent of comparison beneficiaries.

We used beneficiaries' nine-digit ZIP code from the Medicare enrollment data. Since the ADI is measured at the Census block group level, we used the nine-digit ZIP code crosswalk to the ADI from the Neighborhood Atlas, which was constructed to link to Census block groups. We used the 2015 ADI, which is based on socioeconomic variables from the American Community Survey in the previous five years, to ensure consistency over the long demonstration in measurement of living in a disadvantaged area. See https://www.neighborhoodatlas.medicine.wisc.edu/ for more information about how ADI is calculated.

Across all years (two pre-demonstration and eight demonstration years), we were unable to find 1.95 percent of beneficiary nine-digit ZIP codes in the ADI crosswalk (a total of 5,272 nine-digit ZIP codes). For these missing nine-digit ZIP codes, we matched them to the next observable ZIP-level, by creating ADI scores using aggregated data from a higher-level measure of ZIP code (eight, seven, six, or five digits). These measures were simple averages of all the nine-digit ZIP codes contained within that higher-level ZIP code. For example, if an IAH or comparison beneficiary had 30605-9999 as their nine-digit ZIP code and it was not listed in the crosswalk, there may be ZIP codes with the same first eight digits in the crosswalk. Therefore, we assigned this beneficiary to the average ADI score for all ZIP codes with those first eight digits. Of the 1.95 percent of unassigned nine-digit ZIP codes, we were able to assign 37 percent to an eight-digit aggregated ADI score, 13 percent to a seven-digit aggregated ADI score, 6 percent to a six-digit ADI score, and 42 percent to a five-digit aggregated ADI score. This left only 2 percent of the original 1.95 percent that we were unable to assign (or a total of 95 ZIP codes) and had to drop from the regression we estimated using the ADI measure; we did not exclude those observations from any other regression.

One limitation of using the ADI score to measure disadvantage is that, even though the ADI is a composite of 17 variables from the American Community Survey, the way it is constructed puts a large weight on only the area median home value (Hannan et al. 2023). Therefore, there are limitations to interpreting the ADI as a summary measure of deprivation across multiple dimensions. However, even if

the ADI cannot be interpreted as a composite measure as originally intended, it still represents more disadvantaged areas, and the previous research that shows that the ADI is highly predictive of clinical outcomes still holds true.

4.4. 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) ambulatory care, and (4) health outcomes (Exhibit C.15).

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 per beneficiary per month (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 six of 12 months in the demonstration year), the annualized number of hospital admissions would be eight.

Exhibit C.15. Outcome measures used for regression analysis

Measure
Medicare expenditures per beneficiary per month
Total
Inpatient
Home health[a]
Outpatient
Skilled nursing facility
Clinician/supplier
Hospice
Durable medical equipment
Advanced imaging
Other imaging
Testing

Exhibit C.15 (*continued*)

Measure
Hospital use
Number of outpatient ED visits per beneficiary per year
Number of potentially avoidable outpatient ED visits per beneficiary per year (AHRQ PQI)[b]
Number of hospital admissions per beneficiary per year[c]
Number of hospital admissions preceded by an ED visit per beneficiary per year
Number of potentially avoidable hospital admissions 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]
Ambulatory care
Primary care visits
Specialty care visits
Chronic care management services
Health outcomes
Death
Days at home
Entry into a nursing home

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.

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

[c] This includes hospital admissions and observation stays.

[d] Qualifying hospital discharges include discharges for beneficiaries 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. HBPC 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 HBPC 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.

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

Addressing likely fraudulent spending in 2023 (Year 10). Fraudulent billing associated with urinary catheter supplies in Medicare was exposed in 2024 (Office of the Inspector General 2024). We examined changes in payments for intermittent urinary catheter supply codes A4351, A4352, and A4353 in 2022 and 2023 (Years 9 and 10) to determine the scope of the issue for the IAH evaluation. Before 2023, these codes made up less than 4 percent of total DME expenditures for our sample, increasing to more than 40 percent starting in January 2023. Therefore, in 2023 (IAH Year 10), we measured DME and total spending excluding spending related to these three urinary catheter supply codes. We used this

approach because the fraud dollars were so large that if intervention or comparison group was disproportionately affected by it, it could have biased results in 2023. (This decision did not apply to the performance-year analysis we discuss in Appendix D, because that study period ended in 2022.)

Spending on imaging and testing services. We measured spending associated with laboratory and other testing and imaging services in three categories:

1. Total testing (lab and other testing)
2. Advanced imaging (CT, MRI, and nuclear imaging)
3. Standard imaging (x-ray, ultrasound, and miscellaneous imaging)

These measures do not include imaging and testing while a beneficiary was hospitalized.

We identified procedure codes in clinician/supplier and outpatient claims as imaging or testing in three stages to account for the changes in procedure codes during the time of the IAH evaluation. We started with the 2023 Restructured BETOS Classification System (RBCS) which is historical and would provide the most consistent identification. On clinician/supplier claims, any procedure codes that were not flagged by RBCS were identified as testing or imaging based on the BETOS code associated with the procedure codes on the claim (BETOS codes have not been updated since 2016 but would capture any procedure codes not updated in the RBCS for the earlier years of our study). Because BETOS does not exist on outpatient claims, we merged it on by procedure code to identify additional testing or imaging procedure codes. Finally, on both files, any unflagged procedure codes that were in the 70000–79999 range were identified as imaging services and testing services were flagged if the procedure code was in the 80000–89999 range. We further identified advanced imaging services using a subset of RBCS and BETOS values associated with nuclear imaging, CT scans, and MRIs. All other imaging services were considered standard imaging. Expenditures for these three categories of spending on imaging and testing services used the line payments associated with the appropriate procedure codes.

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.

Potentially avoidable outpatient ED visits and hospital admissions. 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 2022: 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.[19]

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.

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 the readmission measure included discharges from nonfederal acute care hospitals for beneficiaries who were enrolled in Medicare FFS, alive at the time of discharge, and not transferred to another acute care facility. The qualifying discharges included beneficiaries discharged to nonacute care settings. Index discharges did not include admissions to Prospective Payment System–exempt cancer hospitals or admissions for beneficiaries without at least 30 days of post-discharge enrollment in FFS Medicare Parts A and B (unless a beneficiary was enrolled in FFS but died within 30 days), beneficiaries discharged against medical advice, primary psychiatric diagnoses, rehabilitation, and medical treatment of cancer or COVID-19.

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. 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 beneficiaries had a qualifying hospital discharge and, if so, whether they had an unplanned readmission within 30 days.

Primary and specialty care visits. These visits reflect those made in ambulatory settings (defined as a non-institutional place of service) with specialty codes identified as primary or specialty care.

In clinician/supplier claims, we measured primary and specialty care visits the following ways:

[19] 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.

- We used procedure codes to limit visits to the following: evaluation and management services; transitional care management services; annual wellness visits; advance care planning; and cognitive assessment and care plan services. Starting in IAH Year 7 (2020), these could be in-person, telehealth, and telephone 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.
- We used specialty codes from the claim to identify clinician specialty. Primary care clinicians were primary care physicians, physician assistants, and nurse practitioners and other advanced practice nurses. Specialty care was limited to physicians (excluding clinicians such as podiatrists and social workers). Because we could not identify the specialty for physician assistants and nurse practitioners, the measure of primary care visits may overcount primary care, and the measure of specialty care visits may undercount specialty care.

In outpatient claims, we measured visits at federally qualified health centers (FQHCs), rural health clinics (RHCs), and critical access hospitals. We identified FQHC and RHC visits using revenue center codes for clinic and home visits and procedure codes for welcome to Medicare or annual wellness visits, non-face-to-face communication, and distant site telehealth services. We also used procedure codes for clinic visits specific to FQHCs, since RHCs do not submit claims with those codes. We identified visits at critical access hospitals as those with a professional fee revenue center code and a procedure code for hospital outpatient clinic visit or any of the procedure codes identified in the previous paragraph. Because FQHCs, RHCs, and critical access hospitals largely provide primary (not specialty) care, we considered all of these visits to be primary care visits.

Entry into a nursing home. The nursing home measure identified beneficiaries who had at least one episode of nursing home care spanning 90 or more days during a given study year. An episode of nursing home 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 nursing home care during a single study year to identify that person as entering nursing home care in that study year. Beneficiaries in nursing home 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 2 of this appendix pose little risk of error for the measure of entry into a nursing home. This is because (1) OASIS assessments are used to identify community days in the Timeline file, and community is the residency status assigned to a beneficiary when data are missing; and (2) CMS has reported that most facilities submitted the MDS within the usual time frame despite the time frame waiver available during the COVID-19 public health emergency in 2020.

Potential episodes of nursing home care began when the beneficiary entered a nursing home (a daily status of S or M) in a given study year. Inpatient days that occurred during a potential episode of nursing home care were considered part of the nursing home episode. An episode of nursing home 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 nursing home 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 nursing home 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.

Days at home. The days at home measure creates a count of the number of days a beneficiary spent at home (as opposed to in a facility) during the demonstration year. According to the Yale documentation for its days at home measure, any day that a beneficiary receives care in one of the following settings is counted as a day in care: inpatient acute (including hospital at home services); inpatient non-acute (inpatient rehabilitation facilities, inpatient psychiatric facilities, and long-term care hospitals); SNF; ED; or observation stay.[20] All other days are considered days at home (also referred to as home days). We used the Timeline file to create daily indicators for whether a beneficiary was in an inpatient or SNF setting during each study year for the evaluation. We used outpatient claims to identify ED and observation stay days. We excluded days in hospice care from facility days, following the Yale approach.

We considered days in a nursing home as home days, following the Yale approach for two reasons. First, once a beneficiary is admitted to an institution, that beneficiary resides in that institution.[21] Second, we wanted this outcome to be conceptually different from the health outcome we use for the evaluation of IAH which measures entry into a nursing home. If we were to count days in a nursing home as facility days, then the days at home outcome measure could be highly correlated with the measure of entry into a nursing home since all beneficiaries in an institution would have no home days for that period. In other words, beneficiaries who entered a nursing home would average many fewer days at home compared with beneficiaries who did not enter such care. Instead, we followed the Yale approach, which counts days in a nursing home as home days.

[20] Yale-New Haven Health Services Corporation – Center for Outcomes Research and Evaluation (CORE). Measure Information Form: Days at Home for Patients with Complex, Chronic Conditions. Information current as of September 2022.

[21] The academic literature has not settled on a recommended approach for measuring days at home for beneficiaries in a nursing home. For example, see van Houtven, C.H., V.A. Smith, N.R. Sperber, C.J. Coffman, and S.N. Hastings. "Advancing the Science of Population-Based Measures of Home-Time." *Healthcare*, vol. 8, number 4, 2020.

5. Methods for estimating impacts

We used a difference-in-differences model to estimate the impact of the opportunity to earn annual performance-based incentive payments 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.

In this chapter, we describe changes in the sample and how those differences affect our estimation. Next, we discuss how the interpretation of estimated effects differs by year due to a different sample of participating practices after Year 5. We then present all methodology related to the analyses.

5.1. Samples used to estimate effects in each demonstration year

Our primary estimates of the effect of the demonstration on outcomes are based on four samples of IAH practices from separate regressions, depending on the years being estimated. The samples are different because practices started to leave the demonstration after Year 5, leading to a different number of participating practices in performance years 6-10. For Years 1 to 5, we estimated effects using a sample containing the full 14 practices that participated through Year 5 of the demonstration. For Year 6, we estimated effects using a separate regression from a sample containing the 12 practices that participated in Year 6. For Year 7, we estimated effects using a separate regression from a sample containing the 10 practices that participated in Year 7. For Years 8 and 9, we estimated effects using a separate regression from a sample containing the seven practices that participated in Years 8 and 9. And finally, for Year 10 we estimating effects using a separate regression containing just practice M, the only participating practice in that year (see Chapter 2 of this appendix for additional information about why we did not retain IAH practices after they stopped participating in the demonstration, and see Appendix B for information about the timing and reasons why IAH practices exited the demonstration).

5.2. Interpretation of estimated effects

In annual reports that focused on Years 1 to 6, as well results from this report focused on Year 10 (Appendix C.9), we interpreted the estimated effects from our difference-in-differences model as estimated effects of the IAH payment incentive (see Exhibit C.1 for more information). Effects from Years 7 to 9, however, should be interpreted differently—specifically, as effects of IAH *during 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 in Years 1 to 6 would have been due to changes made by IAH practices because of the demonstration payment incentive. In Years 7 to 9 (2020 to 2022), the estimates include any changes to the relative effectiveness of care for IAH and comparison beneficiaries if any of the following occurred:

- IAH practices adopted new care delivery strategies during the COVID-19 pandemic.
- Strategies that IAH practices had been using since before the IAH demonstration increased in effectiveness during the pandemic.
- Changes in care experienced by the comparison group did not affect the IAH group.

In addition to potential changes in the relative effectiveness of care, unmeasured factors may have played a larger role in Years 7, 8, and 9 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 Years 7 to 9 should be interpreted as estimated effects of IAH during the COVID-19 pandemic. In addition, there could be an increased risk of confounding in the estimated effects in those years (relative to the risk for Years 1 to 6) 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 Years 7 to 9.

5.3. Difference-in-differences model

5.3.1. 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 nine years of the demonstration, as described in Appendix Exhibit C.1. Equation (1) shows the model we estimated for each outcome in Year 10:

(1)

$$
\begin{aligned}
Y_{it} = {} & \alpha + X_{it}\beta + \tau \bullet treatment_{it} + \gamma_{-1}\, PD_1 + \gamma_1\, DY_1 + \gamma_2\, DY_2 + \gamma_3\, DY_3 + \gamma_4\, DY_4 + \gamma_5\, DY_5 + \gamma_6\, DY_6 \\
& + \gamma_7\, DY_7 + \gamma_8\, DY_8 + \gamma_9\, DY_9 + \gamma_{10}\, DY_{10} + \theta_{-1} treatment_{it} \bullet PD_1 + \theta_1 treatment_{it} \bullet DY_1 + \theta_2 treatment_{it} \bullet DY_2 \\
& + \theta_3 treatment_{it} \bullet DY_3 + \theta_4 treatment_{it} \bullet DY_4 + \theta_5 treatment_{it} \bullet DY_5 + \theta_6 treatment_{it} \bullet DY_6 + \theta_7 treatment_{it} \bullet DY_7 \\
& + \theta_8 treatment_{it} \bullet DY_8 + \theta_9 treatment_{it} \bullet DY_9 + \theta_{10} treatment_{it} \bullet DY_{10} + \varepsilon_{it}
\end{aligned}
$$

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); $DY_1 - DY_{10}$ 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 $\theta_1 - \theta_{10}$, 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 $\theta_1 - \theta_{10}$ 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 $\theta_1 - \theta_{10}$ into difference-in-differences impact estimates using the following steps using the estimated impact in Year 10 as an example:

1. Using the coefficients obtained from Equation (1), we calculated the average outcomes for the IAH treatment and comparison groups in Year 10. We adjusted the yearly average outcomes for both groups to reflect the covariate distribution of the IAH group in each demonstration year. For example, we used the mean covariate values of the Year 10 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 intervention 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 10.
3. We calculated the change in the difference between the IAH and matched comparison groups in Year 10 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 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. Later sections of this chapter describe clustering and weighting in detail.

5.3.2. 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)

$$\begin{aligned}\log(T_{it}) = {} & X_{it}\beta + \tau \bullet treatment_{it} + \gamma_{-1}\, PD_1 + \gamma_1\, DY_1 + \gamma_2\, DY_2 + \gamma_3\, DY_3 + \gamma_4\, DY_4 + \gamma_5\, DY_5 + \gamma_6\, DY_6 + \gamma_7\, DY_7 \\ & + \gamma_8\, DY_8 + \gamma_9\, DY_9 + \gamma_{10}\, DY_{10} + \theta_{-1} treatment_{it} \bullet PD_1 + \theta_1 treatment_{it} \bullet DY_1 + \theta_2 treatment_{it} \bullet DY_2 \\ & + \theta_3 treatment_{it} \bullet DY_3 + \theta_4 treatment_{it} \bullet DY_4 + \theta_5 treatment_{it} \bullet DY_5 + \theta_6 treatment_{it} \bullet DY_6 + \\ & \theta_7 treatment_{it} \bullet DY_7 + \theta_8 treatment_{it} \bullet DY_8 + \theta_9 treatment_{it} \bullet DY_9 + \theta_{10} treatment_{it} \bullet DY_{10} + z_{it}\end{aligned}$$

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; $DY_1 - DY_{10}$ 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 $f(.)$.

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.[22] 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.

[22] 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.

After estimating the survival regression, we transformed $\theta_1 - \theta_{10}$ 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 each year. 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.

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.

5.3.3. 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 a nursing home 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, 20.5 percent of IAH beneficiaries died during Year 8 compared with 27.7 percent of matched comparison beneficiaries).[23] 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 a nursing home) 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 probability of dying in the second

[23] 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.

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)
$$\begin{aligned} P(Y_{it}=1) = {} & \alpha + X_{it}\beta + \tau.treatment_{it} + \beta_h h_{it} + \beta_{Rh} treatment_i h_{it} + \beta_s survdays_{it} + \gamma_{-1}\, PD_1 \\ & + \gamma_1\, DY_1 + \gamma_2\, DY_2 + \gamma_3\, DY_3 + \gamma_4\, DY_4 + \gamma_5\, DY_5 + \gamma_6\, DY_6 + \gamma_7\, DY_7 + \gamma_8\, DY_8 + \gamma_9 DY_9 + \gamma_{10} DY_{10} \\ & + \theta_{-1} treatment_{it} \bullet PD_1 + \theta_1 treatment_{it} \bullet DY_1 + \theta_2 treatment_{it} \bullet DY_2 + \theta_3 treatment_{it} \bullet DY_3 \\ & + \theta_4 treatment_{it} \bullet DY_4 + \theta_5 treatment_{it} \bullet DY_5 + \theta_6 treatment_{it} \bullet DY_6 + \theta_7 treatment_{it} \bullet DY_7 \\ & + \theta_8 treatment_{it} \bullet DY_8 + \theta_9 treatment_{it} \bullet DY_9 + \theta_{10} treatment_{it} \bullet DY_{10} + \omega_{it} \end{aligned}$$

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 $\theta_1 - \theta_{10}$ 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.

5.3.4. 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 beneficiaries 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 HBPC 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.[24] Implementing this approach meant that all IAH beneficiaries in a given practice 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 demonstration). To correctly identify the clustering effect in the IAH group, we did not include practice fixed effects in the regression equation.[25]

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

[24] Accounting for clustering at the practice level for the intervention 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 intervention group and the value of the beneficiary ID for the comparison group.

[25] Ideally, including practice 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 practice (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 practice in a stratified design (Schochet 2008). Relative to the practice 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.

demonstration practices to HBPC 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 practices 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 practices.

5.3.5. 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. 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 9 sample on January 1, 2022, and died on June 20, 2022, 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 practice and year.[26] As we described, we implemented hybrid clustering adjustments but could not use practice fixed effects (an indicator for each practice). Because beneficiaries had different eligibility weights, the number of weighted IAH beneficiaries in a given practice and year might differ from the number of weighted comparison beneficiaries in the same practice and year if we used the composite weight without rescaling it. For this reason, we rescaled the weights for comparison beneficiaries by practice

[26] For more information, refer to the evaluation report covering Year 4 of the IAH demonstration.

and year so that for each year, the weighted number of IAH beneficiaries equaled the weighted number of comparison group beneficiaries for each practice. 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.[27] For other binary outcomes (probability of unplanned readmission and entering a nursing home), 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.

5.3.6. 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 in our regressions: (1) variables describing eligibility for the demonstration; (2) demographic and enrollment characteristics; (3) ADL indicators; and (4) measures of health status, including HCC risk score, HCC indicators, and chronic condition indicators. We included all HCC indicators and categories of HCCs used for matching (Exhibit C.8), but some of the other control variables were included in 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 Year 9, we used the Year 9 values of all control variables whether or not a beneficiary appeared in the sample in an earlier demonstration year. More information about control variables in regressions is available upon request.

5.3.7. 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

[27] 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 practice and year were equal to each other.

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. These results suggested that the parallel-trend assumption held for those outcomes between those two years. In the Year 9 sample, there were a few instances where the difference-in-differences estimate for two years before the demonstration was statistically significant or was large (greater than 5 percent but not statistically significant) and moving in a direction that could lead to substantively different conclusions in our findings. In those cases, we included that information as we described results in the report.[28]

5.3.8. 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 re-estimated the beneficiary-weight regression, leaving out one practice at a time.[29] Specifically, we estimated 16 regressions (treating each member of the Mid-Atlantic Consortium separately), with each regression excluding the IAH beneficiaries from one practice and their matched comparisons in all years. If all regressions showed similar estimates 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 practice strongly influenced the full sample estimate. In Appendix C, we report the estimated effects from each of the 14 regressions for total spending, inpatient spending, primary care visits, and hospital admissions preceded by an ED visit.

[28] 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.

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

5.3.9. Average annual 10-year effect of the demonstration

In addition to estimating the yearly effect of the demonstration, we estimated the average annual effects for all outcomes through 10 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, Year 7, Years 8 and 9, and Year 10. Combining these estimated effects from individual years to create an estimated average effect across the first eight 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 HBPC practices. This is particularly true in this case for two reasons. First, in Years 7, 8, and 9, 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 Section 5.2 of this appendix). If effects were different in those years because of the pandemic, combining those years with pre-pandemic years would create an average estimate that is difficult to interpret. Second, our estimates of the effect of the demonstration on outcomes are based on four samples of IAH practices from separate regressions, as described in Section 5.1 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 HBPC greatly affected estimated effects while it was in the demonstration. Therefore, combining estimates from Years 1 to 5 of the demonstration (14 practices) with estimates from Year 6 (12 practices), Year 7 (10 practices), Years 8 and 9 (seven practices), and Year 10 (one practice) 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 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 practice 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 practices 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 estimated models estimating the yearly effects of IAH for each outcome, saving their coefficients and variance-covariance matrices for each of three models with different samples: (1) Years 1 to 5 with all 14

IAH practices; (2) Year 6 using only 12 practices that participated in Years 1 to 6; (3) Year 7 using only 10 practices that participated in Years 1 to 7; (4) Years 8 and 9 using only seven practices that participated in Years 1 to 9; and (5) Year 10 using only the one practice that participated in Year 10. Those models produce estimates that correspond to the reported effects of IAH in each year (as shown in Appendix C, Chapter 8).[30]

2. We used the Stata command for seemingly unrelated estimation to combine the variance-covariance matrix of the models using cluster robust standard errors.[31,32]
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, Year 7, Years 8 and 9, and Year 10 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 practice at a time in all years, specifically for total spending, inpatient spending, primary care visits, and hospital admissions preceded by and ED visit.

5.3.10. 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 Years 7 to 9 (2020 to 2022) that controlled for COVID-19 diagnosis and COVID-19 hospitalization in our regressions.[33] 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 may cause us to interpret our main results differently. We did not find large differences with and without these controls, suggesting that COVID-19 hospitalizations or diagnosis did not play a direct material role in our results.

5.3.11. Subgroup analysis

To better understand the estimated effects, we estimated effects for subgroups of beneficiaries who may have benefited more from the improvements IAH practices made in care delivery because of the IAH payment incentive. We defined the following subgroups for analyses: (1) beneficiaries needing assistance from another person with more ADLs (five or all six ADLs) versus beneficiaries needing

[30] For estimating effects on readmissions and the probability of entering a nursing home, our primary annual effect estimates are calculated using a bootstrap estimation procedure. Because the steps here to produce average annual effects use multiple samples and regressions estimated simultaneously, it would be computationally complex to also bootstrap this procedure. To estimate average annual effects for these two outcomes, we removed the bootstrap, estimating standard errors the same way we do for other outcomes.

[31] For non-linear outcomes, we report marginal effects, translating model coefficients into interpretable units (for example, percentage points). Stata's built in seemingly unrelated estimation command does not easily allow for estimation of marginal effects. To retain this functionality, we manually replicated the Stata command by stacking the different estimation samples (from Step 1) and running them simultaneously in a fully interacted model with interactions for sample. This produces identical results to the Stata command and allows for more straightforward estimation of marginal effects.

[32] For estimating effects on mortality, our primary annual effect estimates use parametric survival models. These models did not converge using the procedure we used to estimate average annual effects in the non-linear case we described for readmissions and the probability of entering a nursing home. Instead, we used a logistic regression model to estimate the average annual effects across all nine years on mortality.

[33] For more details, refer to the evaluation reports covering Year 7, Year 8, and Year 9 of the demonstration.

assistance with fewer ADLs (two to four); (2) beneficiaries with dual Medicare and Medicaid coverage versus beneficiaries with Medicare alone; (3) 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 (4) beneficiaries residing in a more versus less disadvantaged area as measured by the ADI. For the ADI, we created high and low categories by splitting the continuous default score at the 75th percentile for the nation (high ADI = above 75th percentile, low ADI = below 75th percentile) as described in Chapter 4 of this appendix. In our data, about two thirds of the sample in Year 9 was in less disadvantaged ADI areas and one third of the sample was in more disadvantaged ADI areas. Sample sizes for all subgroups are in Exhibit C.26.

To estimate effects for these subgroups, we ran regressions for total Medicare 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 for one group (for example, beneficiaries with dual Medicare and Medicaid coverage) were different from the other group (for example beneficiaries with Medicare alone). We used a similar approach to estimate 10-year average effects. That is, we started with the 10-year average effect model (which is a series of regressions by participating sample, see Appendix 5.3.9), and interacted binary subgroup indicators with all covariates (including the sample participation indicator). This additional set of interactions allowed for the same tests between groups described above.

6. Methods for accounting for other initiatives and payments

Other programs administrated by CMS that take place concurrently with the IAH demonstration could influence our estimates of the effect of the demonstration 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 the demonstration: the Quality Payment Program and ACOs.

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

6.1.1. 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 9 (2022). In the first three years of issuing payments, payments were small, with a maximum adjustment of just 2.3 percent (CMS 2023c). 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.[34]

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 the 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.

[34] Though adjustments were made based on the year of performance, we account for Medicare expenditures in the year the payments were made.

- 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 beneficiaries 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.

6.1.2. Payments included when estimating impacts of the IAH demonstration

We included MIPS adjustments on payments made to MIPS-eligible clinicians from 2019 to 2023 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 these years 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 the demonstration.

We did not include Advanced APM payments made to clinicians 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 beneficiary in the comparison group likely had many beneficiaries 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 lower than MIPS even despite its growing popularity in recent years. About 386,000 clinicians (of the approximately 1 million nationwide) received an Advanced APM payment in 2022 (Cottrill et al. 2025), and while this number is growing, it still represents a minority of clinicians nationwide. Therefore, any effects of additional payments will likely 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 beneficiary 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. In Years 7-10, clinician/supplier claims accounted for only about 15 percent of total expenditures for both IAH beneficiaries and 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.

6.1.3. 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.[35]

6.1.4. Accounting for population-based payment adjustments in claims

Population-based payment (PBP) adjustments are adjustments made to Medicare claims to offset payments made to providers participating in certain CMS models, like the Primary Care First model or Next Generation ACO model.[36] Providers in those models are given fully or partially prospective capitated payments and their claims are subsequently adjusted, usually downward, to account for these payments.

By default, these adjustments are included in Medicare claims, so without the accompanying payment amount made outside of the claims, the claims record likely undercounts true Medicare spending. In accordance with CMS guidance, we conducted a sensitivity analysis that reverses the PBP adjustments in claims, returning the claim amount to what CMS would have paid without models that offer PBPs.

[35] 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.

[36] According to CMS guidance, the list of models including PBPs during the IAH demonstration (2012 to 2023) are: Pioneer, Next Generation ACO, Comprehensive Primary Care Plus, Vermont All-Payer ACO, ACO REACH, Primary Care First, and Kidney Care Choices.

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

6.2.1. 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 beneficiaries' 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 2022, there were 483 ACOs in the Medicare program with 11.0 million assigned Medicare beneficiaries, up from 220 organizations and 3.2 million beneficiaries at the start of 2013 (CMS 2023a). In 2022, the following ACO initiatives were active: the Global and Professional Direct Contracting Model, the Vermont All-Payer ACO model, and the Medicare Shared Savings Program (SSP). Most of the IAH practices joined an SSP ACO during the demonstration. By Year 9, all but one of the seven remaining IAH practices participated in an ACO (in Year 10, the only remaining practice was in an ACO), leading to 94 percent of IAH beneficiaries participating in an ACO in that year. By comparison, nearly half of comparison beneficiaries participated in an ACO in Year 9.

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 beneficiaries' 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 beneficiaries 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-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.[37]

6.2.2. Assigning ACO participation to IAH and comparison beneficiaries

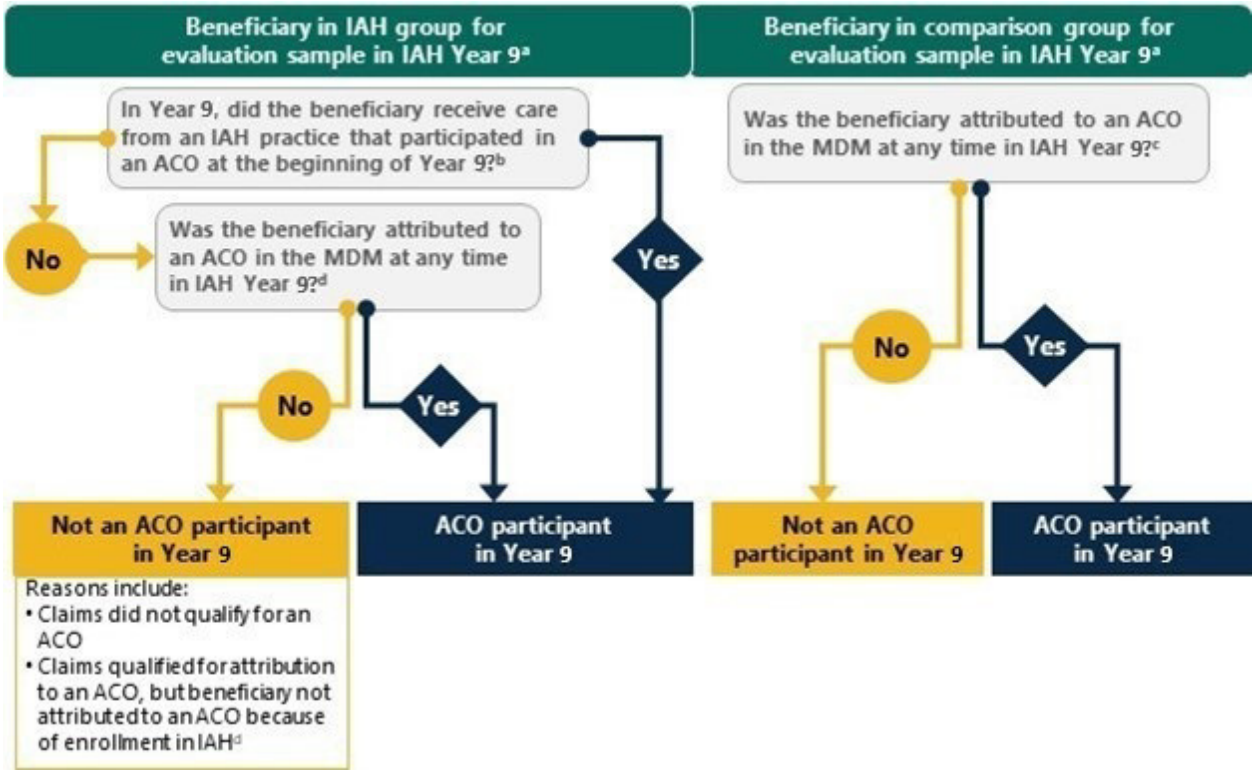
A key challenge to assessing the influence of ACO participation on the estimated impact of IAH is to identify beneficiaries in our sample reached by ACOs. Attribution to an ACO initiative depends on having qualified claims from ACO providers as measured by Medicare claims that have the National

[37] 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 practices. 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.

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.[38] Beneficiaries in the IAH evaluation sample but not enrolled in the IAH demonstration (discussed in Chapter 2 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 practices 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 which they received care in a given demonstration year. A description of this process using Year 9 as an example is shown in Exhibit C.16.

Exhibit C.16. ACO assignment for IAH and comparison group beneficiaries



[38] 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 C.16 (*continued*)

Note: IAH Year 9 was January to December 2022.

[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 attribution to an ACO initiative and participation in other CMS 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.

We identified IAH beneficiaries who participated in an ACO in each study year (demonstration Years 1 to 10 and two pre-demonstration years) based on the following logic:[39]

- We assigned IAH beneficiaries to an ACO if their IAH practice TIN was participating in an ACO at the beginning of the study year.[40]
- 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.

6.2.3. 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 C.17 shows ACO participation for each IAH practice at the beginning of each demonstration year according to information provided by IAH practices. No practices 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 9 (94.0 percent), with the last and only practice participating in an ACO (100 percent) in Year 10 (Exhibit C.18). Among comparison beneficiaries, participation in an ACO increased steadily over time to 48.8 percent in Year 6, dropping slightly to 40.6 in Year 7 but returning to 48.7 in Year 9.

[39] IAH enrollment in the MDM does not align perfectly with inclusion in the IAH group we identified for the evaluation. See Chapter 2 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.

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

Exhibit C.17. IAH practices' ACO participation at the start of each IAH demonstration year

Demonstration practice location	Year 1 Jun 2012–May 2013	Year 2 Jun 2013–May 2014	Year 3 Jun 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	Years 8 and 9 Jan 2021–Dec 2022	Year 10 Jan 2023–Dec 2023
Practice that participated in Year 10									
Long Island, NY					✓	✓	✓	✓	✓
Practices that participated in Years 1 to 9									
Dallas, TX[a]				✓	✓	✓	✓	✓	n.a.
Flint, MI[a]	✓	✓	✓	✓	✓	✓	✓	✓	n.a.
Jacksonville, FL[a]				✓	✓	✓	✓	✓	n.a.
Lansing, MI[a]				✓	✓	✓	✓	✓	n.a.
Milwaukee, WI[a]				✓	✓	✓	✓	✓	n.a.
Philadelphia, PA[b]									n.a.
Richmond, VA[b]									n.a.
Baltimore, MD/ Washington, DC[b]				✓	✓	✓			n.a.
Practices that left the demonstration after Year 5									
Austin, TX						n.a.	n.a.	n.a.	n.a.
Cleveland, OH				✓	✓	n.a.	n.a.	n.a.	n.a.
Practices that left the demonstration after Year 6									
Boston, MA				✓	✓	✓	n.a.	n.a.	n.a.
Durham, NC							n.a.	n.a.	n.a.
Practices that left the demonstration after Year 7									
Brooklyn, NY				✓	✓			n.a.	n.a.
Portland, OR								n.a.	n.a.
Wilmington, DE				✓	✓	✓	✓	n.a.	n.a.

Source: CMS's MDM database and information provided by IAH practices 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.23, 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.

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

Exhibit C.18. Number and percentage of IAH and comparison beneficiaries assigned to an ACO, by evaluation year

	Second pre-demon-stration year: Jun 2011–May 2012[a]	Year 1 Jun 2012–May 2013[a]	Year 2 Jun 2013–May 2014[a]	Year 3 Jun 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]	Year 8 Jan 2021–Dec 2021[b]	Year 9 Jan 2022–Dec 2022[b]	Year 10 Jan 2023–Dec 2023[b]
Number of IAH beneficiaries	107	1,773	1,392	1,498	7,001	7,445	5,509	4,153	3,159	2,462	167
Percentage of IAH beneficiaries	1.5	21.6	19.2	19.8	73.7	74.8	68.9	82.1	94.5	94.0	100.0
Number of comparison beneficiaries	1,246	5,412	8,431	10,685	15,846	19,670	15,872	8,468	6,129	5,217	305
Percentage of comparison beneficiaries	3.9	16.0	26.1	34.2	41.3	47.5	48.8	40.6	45.3	48.7	39.2

Source: CMS's MDM database and information provided by IAH practices 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 Mid-Atlantic Consortium, Years 1 to 3 and the two pre-demonstration years began in September rather than June.

[b] Year 10 analyses included only the one practice that participated in Year 10. Year 8 and 9 analyses included only the seven practices that participated in those years. 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.

6.2.4. Accounting for ACO participation in regression analysis

To assess whether ACO participation influenced the effect of the demonstration on 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 the demonstration differed depending on whether we controlled for ACO participation. That said, by Year 9, with nearly 95 percent of IAH beneficiaries participating in an ACO (including 100 percent in Year 10), it was difficult to separate effects of the demonstration from the effects of ACOs in later years.

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 beneficiary population or care delivery that affected Medicare expenditures, such changes probably did not occur immediately after joining an ACO.

7. Assessing strength of evidence

In addition to stating quantitative results in the final report, we categorized the strength of the evidence for each result. The purpose of assessing the strength of evidence is to give the reader our overall level of confidence in a result estimated by the regression model. Because we included results from two different analyses in this report—the performance-year analysis and the three-year analysis of beneficiaries new to HBPC—we assessed the strength of evidence for each outcome separately for each analysis.

This section describes the criteria we used to assign strength of evidence throughout the report and summarizes the level of evidence for each outcome.

7.1. Criteria for assessing strength of evidence

7.1.1. Magnitude and statistical significance of results

The first criterion relates to size (or magnitude) and statistical significance of the result. We considered effects to be large if they were greater than or equal to 5 percent, moderate if they were between 3 and 5 percent, and small if they were less than 3 percent. We considered effects to be statistically significant if the *p*-value was less than 0.10. Large effects that were statistically significant represented the strongest evidence.

We used the first criterion to identify results that required further assessment using the remaining criteria and to help define the overall level of evidence. We took this approach for two reasons. First, some of the remaining criteria did not apply when we saw no evidence of an effect (that is, a result around 0 or 1 percent). For example, we did not consider whether the result was in the same direction as the hypothesized relationship between IAH and that outcome when the result was small and not statistically significant. Second, when determining the overall level of evidence, we wanted to distinguish between (1) a strong level of evidence that IAH did not affect an outcome (no evidence of an effect) and (2) a weak level of evidence that IAH decreased or increased an outcome (see Appendix C.7.2 below for more details).

7.1.2. Clear mechanism of effect in hypothesized direction

This criterion assesses whether findings from the evaluation had a clear mechanism of action. In Appendix Exhibit B.8, we list hypotheses and mechanisms about how we expected HBPC from an IAH practice and the IAH payment incentive to affect outcomes. These expectations are based on our interviews with IAH administrators, clinicians, and staff; our understanding of published literature about HBPC; and the IAH legislation. We considered the evidence stronger if estimated (non-zero) effects aligned with these hypotheses. For example, we hypothesized that HBPC with the payment incentive would lead to fewer hospital admissions (not more).

7.1.3. Alignment between direction of results and qualitative evidence

For this criterion, we assessed whether the estimated effects aligned with qualitative evidence we collected during the evaluation. When quantitative effects aligned with qualitative findings, the level of evidence was stronger. Although we did not have direct qualitative evidence to directly support (or

refute) hypotheses for each outcome, we used information from interviews and beneficiary surveys to determine whether we expected the demonstration to have affected outcomes. For example, respondents at IAH practices reported making only modest changes to care delivery in response to the opportunity to earn incentive payments, which suggests effects of the incentive payments on outcomes like ED visits and hospital admissions might have been small.

7.1.4. Consistency across time or with other analyses

This criterion assessed the consistency of effects across time (effects in multiple years) and across sensitivity analyses within a study (such as sensitivity analyses excluding one IAH practice at a time or excluding very high [outlier] spending). Results with more consistency had stronger levels of evidence because consistency lends credibility that a result is not due to random variation in a single year or to a specific methodological decision. Because the demonstration was small, we aimed to avoid highlighting single-year effects or effects that were not supported by sensitivity analyses. We looked for similarity of direction (increase or decrease) and size (large, moderate, and small) across years. Consistency does not mean all results needed to be exactly (or almost exactly) the same across time or across analyses. It is instead meant to avoid overinterpretation of results in isolation that could be driven by statistical noise (randomness) given the small size of the demonstration.

7.1.5. Study design and assessment of assumptions

This criterion assessed the strength of the study design and the assumptions needed to interpret effects as causal. For nearly all analyses and outcomes, we used a strong difference-in-differences design (see Appendix C.5). One analysis—the examination of effects on death in the three-year analysis—could not use that design because the outcome was not observed in the pre-period. For that analysis, we used a simpler cross-sectional design that compared IAH beneficiaries to comparison beneficiaries in the intervention period only (see Appendix D). Though results from this analysis are still important evidence for the report, we categorized it as lower strength of evidence than the rest of the findings generated from a stronger study design.

For analyses that used a strong difference-in-differences design (all except death in the three-year analysis), we assessed a key statistical assumption required to interpret effects as being caused by the demonstration or starting HBPC (that is, causal effects), as opposed to simply being associated with them. Specifically, the difference-in-differences design assumes that the comparison group represents a good counterfactual—what *would have* happened in the intervention group absent the intervention. This assumption is not directly testable because we cannot observe what would have happened absent IAH. We can, however, observe outcome trends in the intervention and comparison groups before the demonstration began (pre-period). If those pre-period trends are similar (ideally nearly identical, but at least parallel), the more confidence we have that comparison group does indeed represent a good counterfactual.

For each outcome and analysis, we assessed whether the *difference* in pre-period trends between IAH and comparison beneficiaries was: (a) statistically different using $p < 0.10$; (b) large (greater than 5 percent); or (c) in a direction that could explain results (undermining effects we attribute to IAH), had those trends continued into the demonstration period (Exhibit C.19), as we describe here:

- The most concerning case, and therefore the weakest in terms of strength of evidence, is the case where all three factors we described as (a), (b), and (c) occurred. This case would most likely suggest a violation of the parallel trends assumption and would weaken our conclusions about the results. We saw these three factors occur together in the performance-year analysis for hospice spending, entering a nursing home, testing spending, and primary care visits. We saw these three factors occur together in the three-year HBPC analysis for outpatient spending, hospice spending, and specialty care visits.
- It is still a serious reduction in the strength of evidence when a direction in pre-period trends that could explain the results (c) occurs with either only one of (a) or (b); however, the likelihood of the violation could be less.
- Least serious is when (a) or (b) occur without (c); indicating a violation of the parallel trends assumption that would not directly change our conclusion.

Exhibit C.19. Assessment of parallel trends assumptions by outcome and analysis

Outcome	Trends for the IAH and comparison groups were statistically different ($p < 0.10$) before IAH began (a)	Difference in trends for the IAH and comparison groups was large (>5 percent) before IAH began (b)	Difference in trends for the IAH and comparison groups were in a direction that could explain results during the demonstration, had they continued (c)
Performance-year analysis[a]			
Total Medicare spending			
Inpatient spending			
Home health spending			
Outpatient spending		X	X
SNF spending			
Clinician/supplier spending			
Hospice spending	X	X	X[d]
DME spending			
Hospital admissions			
Outpatient ED visits			
Hospital admissions that were preceded by an ED visit			
Potentially avoidable hospital admissions	X	X	
Potentially avoidable outpatient ED visits		X	
Readmissions			
Entering a nursing home	X	X	X
Death	X	X	
Days at home			

Exhibit C.19 (*continued*)

Outcome	Trends for the IAH and comparison groups were statistically different ($p < 0.10$) before IAH began (a)	Difference in trends for the IAH and comparison groups was large (>5 percent) before IAH began (b)	Difference in trends for the IAH and comparison groups were in a direction that could explain results during the demonstration, had they continued (c)
Advanced imaging spending			
Other imaging spending			
Testing spending	X	X	X[d]
Primary care visits	X	X	X
Specialty care visits		X	X
Three-year analysis[b]			
Total Medicare spending			
Inpatient spending			
Home health spending			
Outpatient spending	X	X	X[d]
SNF spending			
Clinician/supplier spending			
Hospice spending	X	X	X[e]
DME spending	X	X	
Hospital admissions			
Outpatient ED visits	X	X	
Hospital admissions that were preceded by an ED visit			
Potentially avoidable hospital admissions	X		X
Potentially avoidable outpatient ED visits	X	X	
Readmissions			
Entering a nursing home	X	X	
Death[c]	n.a.	n.a.	n.a.
Days at home	X		
Advanced imaging spending	X		X
Other imaging spending	X	X	
Testing spending	X	X	
Primary care visits	X	X	
Specialty care visits	X	X	X

[a] A cell is marked with an X if it meets the criterion in the column based on two years of pre-demonstration data for two or more IAH performance years for the performance-year analysis. For example, if the data for the 12 practices in Year 6 and the 10 practices in Year 7 (but not for practices in other years) indicated a statistically significant difference in trends for IAH and comparison groups before the demonstration began for a particular outcome, then table reflects this with an X for that outcome for the performance-year analysis. We included Year 10 in this assessment, which only had one

Exhibit C.19 (*continued*)

participant. However, including Year 10 did not change any of the cells in the exhibit because there were no outcomes where there were only two years that the criterion occurred and one of those two years was Year 10.

[b] A cell is marked with an X if it meets the criterion in the column based on two years of pre-period data before the index date for the three-year analysis. For example, the trend for entering a nursing home for IAH and beneficiaries from two years prior to one year prior to the index date was a much sharper increase than the trend for comparison beneficiaries (Exhibit D.28). The difference between the two trends was statistically significant ($p < 0.10$). The table reflects this with an X for that outcome for the three-year analysis.

[c] The three-year analysis follows beneficiaries over time, so no beneficiaries died before joining the sample. For this reason, we cannot assess pre-intervention trends in death.

[d] In these three cases, the estimated effect was small, usually not statistically significant, and the outcome was not a primary focus of the evaluation. Therefore, we did not address the parallel trend concern in the main report for these three cases. For example, in the performance-year analysis, the estimated effect of the demonstration was an increase of 2.0% in hospice spending, which was not statistically significant. This small effect is one of several reasons why evidence of this effect was weak, so we did not highlight the result or its pre-period trend in the report. In other cases where an outcome has an X for all three criteria in this exhibit, we address the concern about parallel trends in the pre-demonstration period for that outcome in the report. For example, in Chapter 2 of the report, we note a concern about the pre-demonstration trend for primary care visits in the performance-year analysis and explain how that concern affected our interpretation of the result for that outcome.

[e] For hospice spending in the three-year analysis, there was an X for all three criteria. The difference in the pre-demonstration trend was -$28 PBPM. If we assume that the pre-demonstration trend would have continued throughout the three-year period following the index date, the IAH group would have increased its spending on hospice use relative to the comparison group by an additional $28 PBPM during each of the three follow up years—a total of $84 PBPM by the third year. Even if this had happened, starting HBPC during the demonstration still would have had a large, statistically significant effect on hospice spending because the estimated effects ranged from $150 PBPM in the first year to $203 PBPM in the second year. Because examining the pre-period trend did not change our conclusion about this result, we did not mention the trend in the report.

DME = durable medical equipment; ED = emergency department; SNF = skilled nursing facility.

7.2. How we summarized criteria into a single level of evidence

The five criteria described above were the following:

- Magnitude and statistical significance of results
- Clear mechanism of effect in hypothesized direction
- Alignment between direction of results and qualitative evidence
- Consistency across time or with other analyses
- Study design and assessment of assumptions

Using the criteria described above, we categorized each result in the report by the overall level of evidence. We categorized results separately for the performance-year and three-year analyses. The levels of evidence were:

- **IAH increased/decreased.** Strong level of evidence that IAH reduced or increased the outcome. Strong evidence across all or nearly all criteria.
- **IAH may have increased/decreased.** Moderate level of evidence that IAH reduced or increased outcome. Some stronger and some weaker evidence across the five criteria.

- **IAH probably did not increase/decrease.** Low level of evidence that IAH reduced or increased outcome. Mostly weak evidence across the five criteria.
- **IAH did not increase/decrease.** Strong level of evidence that IAH did not affect outcome. Strong evidence of no effect across all or nearly all criteria.

7.3. Levels of evidence by study and outcome

In Exhibit C.20 we show the summary level of evidence for each outcome for both the performance-year and HBPC analyses.

Exhibit C.20. Strength of evidence by outcome and analysis

Outcome	Strength of evidence: Performance-year analysis	Strength of evidence: HBPC analysis[a]
Total Medicare spending	May have reduced	Reduced
Inpatient spending	May have reduced	Reduced
Home health spending	Did not affect	Increased
Outpatient spending	May have reduced	Reduced
SNF spending	Did not affect	Reduced
Clinician/supplier spending	May have reduced	Did not affect
Hospice spending	Did not affect	Increased
DME spending	Reduced	Increased
Hospital admissions	Probably did not reduce	Reduced
Outpatient ED visits	Did not affect	Reduced
Hospital admissions that were preceded by an ED visit	May have reduced	Reduced
Potentially avoidable hospital admissions	May have reduced	May have reduced
Potentially avoidable outpatient ED visits	Did not affect	Reduced
Readmissions	May have reduced	Reduced
Entering a nursing home	Probably did not increase	Decreased
Death	May have reduced	Did not affect
Days at home	Did not affect	Increased
Advanced imaging spending	Did not affect	Probably did not reduce
Other imaging spending	Reduced	Increased
Testing spending	Did not affect	Increased
Primary care visits	Probably did not increase	Increased
Specialty care visits	May have decreased	Probably did not decrease
Chronic care management services	NA	Increased

[a] Strength of evidence assessment for the three-year analysis is limited to the first year of the study.
DME = durable medical equipment; ED = emergency department; NA = not available; SNF = skilled nursing facility.

8. How evaluation results differed from incentive payment results

8.1. Approach for calculating incentive payments

Practices that participated in the IAH demonstration were eligible to receive annual incentive payments. Specifically, practices could earn incentive payments if their beneficiaries' Medicare expenditures were below the practice's target expenditures and the practice met performance thresholds for a set of quality measures. Calculation of incentive payments was based on (1) comparison of costs incurred (that is, actual expenditures of enrolled IAH beneficiaries during the performance year) with the target expenditures and (2) performance on payment-related quality measures. Actual expenditures of enrolled IAH beneficiaries were determined from Medicare FFS claims data. Target expenditures represented the expected expenditures based on Medicare FFS beneficiaries not exposed to the demonstration. Performance on payment-related quality measures was determined from information IAH practices report (for three of the six measures) and Medicare FFS claims data (for the remaining three measures).[41] For more information about IAH enrollees, see Chapter 2 of this appendix.

Although procedures for measuring performance on payment-related quality measures remained unchanged throughout the IAH demonstration, calculations of actual and target expenditures 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 2023b). For Year 5, calculations for 10 practices used the revised regression-based methodology, and calculations for four practices used the original actuarial methodology.[42]

Beginning with Year 6, 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 were 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-applicable population in each practice" (CMS 2023b). The utilization factor was an add-on to the base risk score that "reflects the level of risk that is 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 2023b). 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 2021). The utilization factor led to an increase in target expenditures; all else equal, higher target expenditures

[41] See Exhibit 5.1 in the report for a list of practice-reported and claims-based measures.

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

resulted in a larger incentive payment. This implicitly assumed that IAH enrollees who were hospitalized and had post-acute care in one year were 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 have overestimated 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 2023b). Trending expenditures forward to the performance year was necessary because the county-level data used to set target expenditures were lagged by about 15 months. For example, for Year 6, CMS used calendar year 2018 expenditures, which were published in April 2020. Trend factors came from the most recently available Medicare Trustees Report and were updated annually.

Equation (4) shows a simplified version of how CMS calculated the spending target for each beneficiary enrolled in the demonstration in Years 6 to 10. Equation (5) shows how CMS calculated the savings or loss PBPM for each practice.

(4) $$\begin{aligned} &target\ expenditures\ =\ average\ monthly\ FFS\ county\ cost * cost\ trend * \\ &\left(beneficiary\ risk\ score + frailty\ factor + utilization\ factor\right) \end{aligned}$$

(5) $$\begin{aligned} &savings\ /\ loss\ PBPM\ = \\ &\qquad average\ target\ expenditures\ PBPM - average\ actual\ expenditures\ PBPM \end{aligned}$$

The maximum incentive payment for each practice in a given year depended on the following factors (CMS 2021):

- 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 2021).
- The number of months of enrollment in the demonstration by IAH 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 was that a practice must have met the performance threshold for the quality measures—payment did 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 had 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 have been $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.

8.2. Differences between evaluation results and incentive payments

8.2.1. Purposes of the evaluation and incentive payment calculation

The evaluation and the incentive payment calculation served different purposes. The evaluation needed 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 required 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 required using a comparison group of beneficiaries that were similar to IAH beneficiaries but did not receive HBPC and—as we describe in the following section—examining changes in expenditures for the comparison group relative to changes for beneficiaries of IAH practices over the same time period. By contrast, the incentive payment calculation needed to offer a more timely way to measure a target spending level to reward participation and performance. As described in the previous section of this chapter, spending targets were set by trending historical expenditures for Medicare FFS beneficiaries and adjusting those expenditures to reflect IAH beneficiaries' health status and past utilization to compare against actual spending to calculate incentive payments.

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

In all years of the demonstration, large differences existed between the total incentive payments paid to IAH practices and the estimated aggregate effect of the demonstration calculated by the evaluation (Exhibit C.21). The evaluation and incentive payment calculation results differed each year because of (1) differences in the sample of beneficiaries included (Exhibit C.2) and (2) differences in the methods used. As we explain in this section, our use of a difference-in-differences approach accounted for much of the discrepancy between evaluation findings and incentive payment calculation results.

Exhibit C.21. Effects of the demonstration on Medicare spending compared with incentive payments

Performance year	Number of IAH practices	Aggregate effect of the demonstration on Medicare spending, in millions	Aggregate effect of the demonstration on Medicare spending, 90% confidence interval	Incentive payments to IAH practices (in millions)
Year 1	14	-$9.4	-$22.1, $3.2	$11.7
Year 2	14	-$2.2	-$18.1, $13.8	$5.3
Year 3	14	-$12.9	-$31.6, $5.9	$7.2
Year 4	14	-$25.4	-$55.9, $5.0	$8.1
Year 5	14	-$31.4*	-$59.8, -$2.9	$6.9
Year 6	12	-$3.2	-$19.2, $12.9	$11.1
Year 7	10	-$22.6**	-$39.3, -$6.0	$18.5
Year 8	7	-$9.9	-$24.7, $4.8	$12.5
Year 9	7	-$8.0	-$20.5, $4.5	$9.7
Year 10	1	-$0.3	-$1.4, $1.9	$5.1
Total across all years		-$124.7	-$257.8, $8.5	$96.0

Source: IAH demonstration data and Medicare claims, 2010–2023.

Notes: All numbers in this exhibit are in millions. Negative numbers denote spending decreases; positive numbers denote spending increases. Aggregate effect calculations before accounting for incentive payments are based on the beneficiary-level estimates shown in Appendix Exhibit C.23 and the number of IAH beneficiary-months in each year.

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

Mathematica used a quasi-experimental difference-in-differences design, which we describe in detail in Appendix C. 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 (6) shows a simplified version of how the evaluation calculated the effect of the demonstration on expenditures. The terms IAH_{yX} and C_{yX} reflect regression-adjusted PBPM expenditures in Year X (such as Year 5) 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.

(6) $$effect\ on\ expenditures = (IAH_{yX} - C_{yX}) - (IAH_{pd} - C_{pd})$$

The difference in spending between the IAH and comparison groups in a given demonstration year is conceptually similar to the difference between actual and target expenditures used in the incentive payment calculation. In other words, $IAH_{yX} - C_{yX}$ in Equation (6) is conceptually similar to the difference between actual and target expenditures used in the incentive payment calculation as shown in Equation (5).

The difference between IAH beneficiaries' average actual expenditures and average target expenditures shown in Equation (5) is not a true measure of savings produced by IAH because the approach used to calculate target expenditures was based on actuarial predictions. It was not designed to account for factors that could not be observed in available data but may have differed between IAH and other beneficiaries *and* affected expenditures. Examples include physical or financial barriers the beneficiary faced when leaving home to receive health care and the beneficiary's preference for a more palliative or more aggressive style of health care. By contrast, the methodology used by the evaluation was designed to account for unobserved factors that differed between IAH and other beneficiaries and affected expenditures.

Some IAH practices stated that it was incorrect or unfair that the evaluation measured the effect of the demonstration on Medicare spending by comparing (1) the change in outcomes for IAH-eligible beneficiaries receiving care from IAH practices before and after the start of the demonstration relative to (2) the change during the same period for a matched comparison group. They felt that this methodology undercounted cost savings to the Medicare program, believing that the HBPC they delivered had been reducing Medicare spending even before the demonstration created a payment incentive. We can infer from results of the three-year analysis that before the demonstration began, IAH practices probably did reduce spending for beneficiaries who started receiving HBPC.

However, the spending reduction caused by IAH practices' general approach to HBPC alone was not relevant for calculating savings to the Medicare program caused by IAH for all IAH beneficiaries. IAH included a service delivery model (HBPC from practices that specialized in HBPC) and an incentive to improve practice performance (payment incentive). The service delivery model alone was in effect before and during the demonstration, and we found no evidence that the demonstration caused IAH practices to substantively change the way they delivered HBPC. Only the payment incentive had the potential to generate spending reductions for the Medicare program that *would not have been achieved without the IAH demonstration*. Thus, the evaluation's performance-year analysis produced the estimated effect under the demonstration (the IAH payment incentive) on Medicare spending.

9. Supplementary tables

Exhibit C.22. Baseline unadjusted means of outcomes

Outcome name	Practices that participated in Years 1–5	Practices that participated in Year 6	Practices that participated in Year 7	Practices that participated in Years 8 and 9	Practice that participated in Year 10
PBPM Medicare spending					
Total	$4,397	$4,186	$4,287	$4,279	$3,976
Inpatient	$1,741	$1,627	$1,690	$1,672	$1,607
Skilled nursing facility	$605	$621	$609	$599	$747
Home health (Part A and B)	$781	$717	$738	$749	$620
Hospice	$153	$162	$162	$205	$151
Outpatient	$253	$247	$246	$252	$106
Clinician/supplier	$715	$675	$697	$705	$592
Durable medical equipment	$150	$136	$145	$142	$152
Service use outcomes					
Number of hospital admissions per 1,000 beneficiaries per year[a]	1,778	1,727	1,767	1,790	1,412
Number of potentially avoidable hospital admissions per 1,000 beneficiaries per year[b]	457	439	450	451	368
Probability of having a qualifying hospital discharge and unplanned readmission within 30 days of discharge	16.0	15.9	19.7	19.7	17.2
Number of outpatient ED visits per 1,000 beneficiaries per year[c]	1,458	1,440	1,413	1,457	680
Number of hospital admissions preceded by an ED visit per 1,000 beneficiaries per year	1,443	1,422	1,447	1,455	1,234
Number of potentially avoidable outpatient ED visits per beneficiary per year[d]	191	185	187	191	121
Health outcomes					
Probability of dying in the demonstration year	14.9	14.9	16.8	16.2	24.1
Probability of entering institutional long-term care in the demonstration year	7.0	8.6	8.7	8.6	8.3

Exhibit C.22 (*continued*)

Outcome name	Practices that participated in Years 1–5	Practices that participated in Year 6	Practices that participated in Year 7	Practices that participated in Years 8 and 9	Practice that participated in Year 10
Percentage of days spent at home	90.1%	90.1%	91.5%	90.7%	91.7%

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

Notes: We calculated the baseline means of outcomes using the sample of IAH-eligible beneficiaries in the year before IAH started.

[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 visits not resulting in a hospital admission, including visits 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 C.23. Average annual 10-year effects

Outcome	IAH	Comparison	Difference	Estimated average annual percent effect across all 10 years[a]	Estimated average annual effect across all 10 years (SE)
Total Medicare spending PBPM	$4,492	$5,590	-$1,098	-4.9%	-$211 ($137)
Inpatient spending PBPM	$1,734	$2,103	-$369	-7.1%	-$122 ($90)
Home health spending PBPM	$714	$485	$229	0.3%	-$2* ($28)
Outpatient spending PBPM	$292	$376	-$84	-11.2%	-$28*** ($10)
SNF spending PBPM	$651	$1,137	-$486	-1.8%	-11 ($31)
Clinician/supplier spending PBPM	$727	$782	-$55	-4.3%	-$30 ($32)
Hospice spending PBPM	$182	$144	$38	2.0%	$3 ($8)
DME spending PBPM	$102	$84	$18	-17.2%	-25*** ($7)
Advanced imaging spending PBPM	$257	$321	-$64	-1.1%	-$3 ($11)
Other imaging spending PBPM	$467	$400	$67	-10.0%	-$61** ($30)
Testing spending PBPM	$696	$525	$171	-2.9%	-$24** ($35)
Total testing and imaging spending PBPM	$1,420	$1,246	$175	-5.1%	-$88 ($57)
Primary care visits	9.5	4.5	5.0	7.4%	0.7 (0.5)
Specialty care visits	3.8	5.0	-1.2	-9.3%	-0.4 (0.3)
Hospital admissions per 1,000 beneficiaries per year	1,730	2,257	-527	-2.3%	-41 (48)
Hospital admission preceded by an ED visit per 1,000 beneficiaries per year	1,403	1,780	-376	-5.8%	-83* (44)
Outpatient ED visits per 1,000 beneficiaries per year	1,566	1,630	-64	-1.7%	-24 (56)
Potentially avoidable hospital admissions per 1,000 beneficiaries per year	399	492	-93	-3.9%	-18 (19)
Potentially avoidable ED visits per 1,000 beneficiaries per year	214	224	-10	1.7%	3 (8)
Death (percentage point change)	214	224	-10	-3.8%	-0.56 (0.59)

Exhibit C.23 (*continued*)

Outcome	IAH	Comparison	Difference	Estimated average annual percent effect across all 10 years[a]	Estimated average annual effect across all 10 years (SE)
Unplanned readmissions within 30 days of discharge among qualifying hospital discharges (percentage point change)	14.57	18.59	-4.02	-4.3%	-0.69 (0.76)
Entering institutional long-term care (percentage point change)	14.65	16.41	-1.76	9.3%	0.65* (0.35)
Percentage of days spent at home	5.4	11.04	-5.63	0.1%	0.11 (0.26)

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

Notes: This exhibit reports the regression-adjusted means of the IAH and matched comparison groups. The estimated average effect across all ten years is a weighted average of the estimates by year shown in Appendix Exhibits C.28–C.35. Standard errors are in parentheses. We computed coefficients and standard errors by using the weighted sample size, which considers the matching and eligibility weights. We obtained regression-adjusted means by applying the estimated regression coefficients to the covariates of IAH beneficiaries in the intervention period. Because of rounding, a difference-in-differences estimate displayed as zero might be shown alongside a percentage effect that exceeds zero.

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

[a] We used the unadjusted IAH group mean in the year before the demonstration to calculate the percentage effect for each demonstration year. Exhibit C.22 reports the baseline unadjusted IAH group mean for all outcomes.

DME = durable medical equipment; ED = emergency department; PBPM = per beneficiary per month; SE = standard error; SNF = skilled nursing facility.

Exhibit C.24. Average annual 10-year effects, dropping one practice at a time

Excluding site	IAH baseline mean (unadjusted)	Estimated average annual percent effect across all 10 years[a]	Estimated average annual effect across all 10 years (SE)
Effects on total Medicare spending			
Full sample	**$4,349**	**-4.9%**	**-$211 ($137)**
A	$4,423	-6.9%	-$307 ($131)**
B	$4,324	-6.3%	-$273 ($156)*
C	$4,448	-5.2%	-$232 ($147)
D	$4,313	-4.9%	-$211 ($125)*
E	$4,341	-5.0%	-$216 ($138)
F	$4,331	-4.8%	-$208 ($140)
G	$4,375	-4.8%	-$208 ($144)
H	$4,339	-4.9%	-$213 ($138)
I	$4,356	-5.0%	-$217 ($140)
J	$4,386	-4.9%	-$215 ($139)
K	$4,336	-4.8%	-$210 ($140)
L	$4,344	-5.0%	-$218 ($140)
M	$4,361	-4.9%	-$212 ($140)
N	$4,300	-4.0%	-$170 ($140)
O	$4,356	-4.3%	-$189 ($145)
P	$4,227	-1.7%	-$73 ($128)
Effects on inpatient spending			
Full sample	**$1,720**	**-7.1%**	**-$122 ($90)**
A	$1,784	-1.1%	-$198 ($77)***
B	$1,693	-9.3%	-$159 ($104)
C	$1,782	-6.7%	-$119 ($93)
D	$1,691	-7.8%	-$132 ($83)
E	$1,719	-7.4%	-$128 ($91)
F	$1,707	-7.0%	-$120 ($92)
G	$1,721	-6.7%	-$116 ($93)
H	$1,707	-7.0%	-$120 ($91)
I	$1,709	-7.1%	-$122 ($92)
J	$1,727	-7.1%	-$123 ($91)
K	$1,702	-7.0%	-$119 ($91)
L	$1,723	-7.5%	-$130 ($92)
M	$1,715	-7.3%	-$126 ($94)
N	$1,708	-5.8%	-$99 ($97)
O	$1,709	-6.0%	-$103 ($94)
P	$1,643	-1.9%	-$31 ($90)

Exhibit C.24 (*continued*)

Excluding site	IAH baseline mean (unadjusted)	Estimated average annual percent effect across all 10 years[a]	Estimated average annual effect across all 10 years (SE)
Effects on primary care visits			
Full sample	**9.5**	**7.4%**	**0.7 (0.5)**
A	9.5	6.8%	0.7 (0.6)
B	9.5	8.6%	0.8 (0.6)
C	9.2	1.7%	0.2 (0.3)
D	9.5	6.7%	0.6 (0.6)
E	9.5	7.3%	0.7 (0.5)
F	9.5	7.4%	0.7 (0.5)
G	9.5	8.7%	0.8 (0.5)
H	9.5	7.4%	0.7 (0.5)
I	9.4	8.1%	0.8 (0.5)
J	9.5	7.8%	0.7 (0.5)
K	9.6	7.2%	0.7 (0.5)
L	9.5	7.6%	0.7 (0.5)
M	9.6	7.9%	0.8 (0.5)
N	9.5	6.2%	0.6 (0.6)
O	9.5	7.5%	0.7 (0.6)
P	9.3	10.4%	1.0 (0.5)**
Effects on hospital admissions preceded by an ED visit			
Full sample	**1,442**	**-5.8%**	**-83 (44)**
A	1,528	-7.7%	-177 (37)
B	1,394	-6.4%	-89 (50)
C	1,464	-4.4%	-64 (40)
D	1,435	-6.1%	-87 (43)
E	1,443	-5.8%	-85 (44)
F	1,422	-6.0%	-86 (45)
G	1,441	-6.2%	-90 (47)
H	1,441	-5.9%	-85 (44)
I	1,434	-5.7%	-82 (44)
J	1,453	-5.9%	-85 (45)
K	1,442	-5.7%	-82 (44)
L	1,442	-5.9%	-85 (44)
M	1,446	-6.2%	-90 (47)
N	1,397	-5.2%	-73 (44)
O	1,443	-6.2%	-89 (48)
P	1,421	-2.4%	-34 (51)

Exhibit C.24 (*continued*)

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

Notes: The effects in this exhibit were calculated the same as results in Appendix Exhibit C.23, leaving out each IAH practice individually. Standard errors are in parentheses. We computed coefficients and standard errors by using the weighted sample size, which considers the matching and eligibility weights. Because of rounding, a difference-in-differences estimate displayed as zero might be shown alongside a percentage effect that exceeds zero. Excluding Site P, effects on potentially avoidable hospitalizations were -2.1% (10 admissions per 1,000 beneficiaries), and effects on unplanned readmissions were -0.2% (-0.03 percentage points); neither was statistically significant at p=0.10.

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

[a] We used the unadjusted IAH group mean in the year before the demonstration to calculate the percentage effect for each demonstration year, shown in the second column of the exhibit.

ED = emergency department; PBPM = per beneficiary per month; SE = standard error.

Exhibit C.25. Average annual 10-year effects, sensitivity results

Excluding site	IAH baseline mean (unadjusted)	Estimated average annual percent effect across all 10 years	Estimated average annual effect across all 10 years (SE)
Total gross Medicare spending base case estimate	**$4,349**	**4.9%**	**-$211 ($137)**
Total gross Medicare spending, without MIPS payments	$4,352	4.9%	-$211 ($137)
Total gross Medicare spending, controlling for ACO participation	$4,360	4.2%	-$183 ($138)
Total gross Medicare spending, winsorizing at the 99th percentile	$4,294	4.6%	-$196 ($125)

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

Notes: Standard errors are in parentheses. We computed coefficients and standard errors by using the weighted sample size, which considers the matching and eligibility weights. Because of rounding, a difference-in-differences estimate displayed as zero might be shown alongside a percentage effect that exceeds zero.

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

[a] We used the unadjusted IAH group mean in the year before the demonstration to calculate the percentage effect for each demonstration year, shown in the second column of the exhibit.

ACO = accountable care organization; MIPS = Merit-based Incentive Payment System; PBPM = per beneficiary per month; SE = standard error.

Exhibit C.26. Subgroup sample sizes for IAH and comparison beneficiaries, all Years 1 to 10

Subgroup	IAH count (percent)	Comparison count (percent)
Dual eligibility for Medicare and Medicaid		
Non-dually eligible	37,553 (60.9%)	164,656 (64.4%)
Dually eligible	24,135 (39.1%)	90,852 (35.6%)
ADLs		
Required assistance from another person with two to four ADLs	23,775 (38.5%)	101,122 (39.6%)
Required assistance from another person with five or six ADLs	37,913 (61.5%)	154,386 (60.4%)
Dually eligibility for Medicare and Medicaid and ADLs		
Dually eligible and required assistance from another person with five or six ADLs	13,973 (22.7%)	53,210 (20.8%)
All other beneficiaries that are not high needs as defined in the row above	47,715 (77.3%)	202,298 (79.2%)
Original reason for entitlement		
Reason for entitlement was age	40,740 (66.0%)	176,343 (69.0%)
Reason for entitlement was disability, ESRD, or both	20,948 (34.0%)	79,165 (31.0%)
ADI[a]		
Less disadvantaged	43,418 (70.4%)	181,556 (71.1%)
More disadvantaged	18,270 (29.6%)	73,952 (28.9%)

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

Notes: This exhibit reports the sample size for each subgroup across all 10 years for IAH and the matched comparison group and the percentage of the group in each subgroup (for example, percentage of IAH beneficiaries who required assistance from another person with two to four ADLs). As we discuss in Appendix A, all 10 subgroups had good balance between IAH and comparison beneficiaries on variables used in propensity score matching.

[a] Residing in a more disadvantaged area is defined as the beneficiary living in a nine-digit ZIP code with an ADI score at or above the 75th percentile across the United States. Residing in a less disadvantaged area reflects a score below the 75th percentile.

ADI = Area Deprivation Index; ADLs = activities of daily living; ESRD = end-stage renal disease.

Exhibit C.27.a. Average annual 10-year effects by subgroups of beneficiaries based on both dual eligibility for Medicare and Medicaid and needing assistance with most or all activities of daily living

	Dually eligible for Medicare and Medicaid and needs assistance with 5–6 ADLs			All beneficiaries not both dually eligible for Medicare and Medicaid and needing assistance with 5–6 ADLs			
	IAH baseline unadjusted mean	Estimated average annual percent effect across all 10 years	Estimated average annual effect across all 10 years (SE)	IAH baseline unadjusted mean	Estimated average annual percent effect across all 10 years	Estimated average annual effect across all 10 years (SE)	Statistically different across groups
Total gross Medicare spending	$5,936	-12.4%	-$738 ($209)***	$3,952	-1.5%	-$60 ($126)	***
Inpatient spending	$2,812	-17.7%	-$499 ($150)***	$1,442	-1.3%	-$18 (150)	***

Exhibit C.27.b. Average annual 10-year effects by subgroups of beneficiaries based on dual eligibility for Medicare and Medicaid

	Dually eligible for Medicare and Medicaid			Non-dually eligible for Medicare and Medicaid			
	IAH baseline unadjusted mean	Estimated average annual percent effect across all 10 years	Estimated average annual effect across all 10 years (SE)	IAH baseline unadjusted mean	Estimated average annual percent effect across all 10 years	Estimated average annual effect across all 10 years (SE)	Statistically different across groups
Total gross Medicare spending	$4,790	-10.6%	-$499 ($214)**	$4,445	0.4%	-$17 ($108)	***
Inpatient spending	$2,211	-14.5%	-$295 ($142)**	$1,719	0.3%	-$5 ($72)	***

Exhibit C.27.c. Average annual 10-year effects by subgroups of beneficiaries based on needing assistance with most or all activities of daily living

	Needs assistance with 5–6 ADLs			Needs assistance with 2–4 ADLs			Statistically different across groups
	IAH baseline unadjusted mean	Estimated average annual percent effect across all 10 years	Estimated average annual effect across all 10 years (SE)	IAH baseline unadjusted mean	Estimated average annual percent effect across all 10 years	Estimated average annual effect across all 10 years (SE)	
Total gross Medicare spending	$5,280	-6.1%	$310 ($153)**	$3,500	-1.9%	-$66 ($126)*	**
Inpatient spending	$2,252	-10.4%	-$225 ($94)**	$1,298	1.5%	$18 ($87)*	***

Exhibit C.27.d. Average annual 10-year effects by subgroups of beneficiaries based on the area deprivation index

	High ADI			Low ADI			Statistically different across groups
	IAH baseline unadjusted mean	Estimated average annual percent effect across all 10 years	Estimated average annual effect across all 10 years (SE)	IAH baseline unadjusted mean	Estimated average annual percent effect across all 10 years	Estimated average annual effect across all 10 years (SE)	
Total gross Medicare spending	$4,908	-7.5%	-$364 ($197)*	$4,431	-3.2%	-$131 ($137)	
Inpatient spending	$2,120	-9.5%	-$191 ($124)	$1,768	-5.5%	-$87 ($92)	

Exhibit C.27.e. Average annual 10-year effects by subgroups of beneficiaries based on original reason for Medicare entitlement

	Original reason for Medicare entitlement: Not aged			Original reason for Medicare entitlement: Aged			Statistically different across groups
	IAH baseline unadjusted mean	Estimated average annual percent effect across all 10 years	Estimated average annual effect across all 10 years (SE)	IAH baseline unadjusted mean	Estimated average annual percent effect across all 10 years	Estimated average annual effect across all 10 years (SE)	
Total gross Medicare spending	$5,371	-5.2%	-$258 (226)	$4,159	-4.5%	-$183 ($115)	
Inpatient spending	$2,416	-5.9%	-$127 ($157)	$1,590	-7.8%	-$306 ($102)	

Exhibit C.27.f. Average annual 10-year effects by subgroups of beneficiaries based on age

	Age greater than or equal to 85 years			Age less than 85 years			
	IAH baseline unadjusted mean	Estimated average annual percent effect across all 10 years	Estimated average annual effect across all 10 years (SE)	IAH baseline unadjusted mean	Estimated average annual percent effect across all 10 years	Estimated average annual effect across all 10 years (SE)	Statistically different across groups
Total gross Medicare spending	$3,670	-3.5%	-$124 (-$121)	$5,051	-5.5%	-$260 (-$160)	
Inpatient spending	$1,285	-7.4%	-$90 (74)	$2,184	-7.3%	-$142 (106)	

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

Notes: Estimates in these exhibits were calculated using a triple difference model that allowed for testing the difference between groups (last column of the exhibit). Standard errors are in parentheses. We computed coefficients and standard errors by using the weighted sample size, which considers the matching and eligibility weights. Because of rounding, a difference-in-differences estimate displayed as zero might be shown alongside a percentage effect that exceeds zero.

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

[a] We used the unadjusted IAH group mean in the year before the demonstration to calculate the percentage effect for each demonstration year, shown in the second column of the exhibit.

ADI = Area Deprivation Index; ADLs = activities of daily living; SE = standard error.

Exhibit C.28. Estimated effect on total Medicare spending by performance year, Years 1 to 10

	IAH	Comparison	Difference IAH - comparison (SE)	Difference-in-differences estimated effect (SE)	Percentage effect[a]
Year 10 (1 practice)	$7,550	$7,951	-$401 ($547)	$158 ($611)	4.0
Year 9 (7 practices)	$5,596	$6,303	-$707 ($289)	-$322 ($304)	-7.5
Year 8 (7 practices)	$5,475	$6,178	-$703 ($240)	-$320 ($289)	-7.5
Year 7 (10 practices)	$5,002	$5,862	-$860 ($180)	-$459** ($206)	-10.7
Year 6 (12 practices)	$4,725	$5,137	-$412 ($108)	-$41 ($126)	-1.0
Year 5 (14 practices)	$4,526	$4,958	-$432 ($158)	-$330* ($182)	-7.5
Year 4 (14 practices)	$4,301	$4,685	-$384 ($177)	-$282 ($205)	-6.4
Year 3 (14 practices)	$4,481	$4,762	-$280 ($129)	-$178 ($158)	-4.1
Year 2 (14 practices)	$4,639	$4,771	-$133 ($124)	-$31 ($139)	-0.7
Year 1 (14 practices)	$4,658	$4,879	-$221 ($146)	-$119 ($97)	-2.7
One year pre-IAH[b] (14 practices)[c]	$4,794	$4,896	-$102 ($186)	-	-
Two years pre-IAH (14 practices)[c]	$4,972	$5,107	-$135 ($190)	-$33 ($57)	-0.8

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

Notes: This exhibit reports the regression-adjusted means of the IAH and matched comparison groups in each year. Estimates in each year use the sample of sites participating in that 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. We obtained regression-adjusted means by applying the estimated regression coefficients to the covariates of IAH beneficiaries in each demonstration year. Because of rounding, a difference-in-differences estimate displayed as zero might be shown alongside a percentage effect that exceeds zero. Estimated effects in Years 7, 8, and 9 are interpreted as effects of IAH during the first three years of the COVID-19 pandemic.

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

[a] We used the unadjusted IAH group mean in the year before the demonstration to calculate the percentage effect for each demonstration year. Exhibit C.22 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 performance year as the difference in means between IAH and comparison beneficiaries in that year minus the difference in the year before the demonstration.

[c] The pre-IAH spending amounts represent the spending of beneficiaries meeting IAH eligibility criteria for all 14 IAH practices in each pre-IAH year. The difference estimate for each pre-IAH year is not the same for IAH performance years with fewer than 14 practices.

SE = standard error.

Exhibit C.29. Estimated effect on spending categories by performance year, Years 1 to 10

Service type	IAH	Comparison	Difference (IAH - comparison)	Difference-in-differences estimated effect (SE)	Percentage effect[a]
Inpatient					
Year 10 (1 practice)	$3,492	$3,431	$61	$415 ($464)	25.8
Year 9 (7 practices)	$2,446	$2,924	-$478	-$103 ($206)	-6.2
Year 8 (7 practices)	$2,342	$2,878	-$535	-$161 ($201)	-9.6
Year 7 (10 practices)	$2,161	$2,791	-$630	-$245* ($138)	-14.5
Year 6 (12 practices)	$1,876	$2,313	-$436	-$66 ($90)	-4.1
Year 5 (14 practices)	$1,834	$2,245	-$410	-$221* ($120)	-12.7
Year 4 (14 practices)	$1,720	$2,065	-$346	-$156 ($134)	-9.0
Year 3 (14 practices)	$1,838	$2,106	-$269	-$79 ($95)	-4.5
Year 2 (14 practices)	$1,900	$2,130	-$231	-$41 ($77)	-2.4
Year 1 (14 practices)	$1,923	$2,168	-$245	-$55 ($61)	-3.2
One year pre-IAH[b] (14 practices)[c]	$1,943	$2,133	-$190	-	-
Two years pre-IAH (14 practices)[c]	$2,032	$2,233	-$201	-$12 ($49)	-0.7
Skilled nursing facility					
Year 10 (1 practice)	$1,160	$1,834	-$675	-$263 ($203)	-35.2
Year 9 (7 practices)	$639	$967	-$329	-$87 ($62)	-14.5
Year 8 (7 practices)	$659	$951	-$293	-$51 ($45)	-8.6
Year 7 (10 practices)	$663	$968	-$305	-$63 ($54)	-10.4
Year 6 (12 practices)	$687	$882	-$196	$36 ($35)	5.8
Year 5 (14 practices)	$693	$900	-$207	-$2 ($46)	-0.3
Year 4 (14 practices)	$676	$900	-$224	-$18 ($50)	-3.0
Year 3 (14 practices)	$705	$911	-$205	$0 ($27)	0.0
Year 2 (14 practices)	$678	$871	-$192	$13 ($25)	2.2
Year 1 (14 practices)	$673	$892	-$219	-$14 ($19)	-2.2
One year pre-IAH[b] (14 practices)[c]	$687	$893	-$206	-	-
Two years pre-IAH (14 practices)[c]	$734	$949	-$215	-$9 ($18)	-1.5
Home health					
Year 10 (1 practice)	$1,026	$760	$266	$42 ($84)	6.8
Year 9 (7 practices)	$898	$620	$277	$78* ($44)	10.4
Year 8 (7 practices)	$841	$605	$236	$37 ($52)	4.9
Year 7 (10 practices)	$742	$532	$210	$13 ($30)	1.8
Year 6 (12 practices)	$662	$475	$187	-$9 ($38)	-1.2

Exhibit C.29 (*continued*)

Service type	IAH	Comparison	Difference (IAH - comparison)	Difference-in-differences estimated effect (SE)	Percentage effect[a]
Year 5 (14 practices)	$703	$494	$208	$4 ($36)	0.5
Year 4 (14 practices)	$660	$473	$187	-$17 ($36)	-2.2
Year 3 (14 practices)	$659	$487	$172	-$33 ($38)	-4.2
Year 2 (14 practices)	$788	$554	$234	$30 ($24)	3.8
Year 1 (14 practices)	$748	$552	$196	-$8 ($17)	-1.1
One year pre-IAH[b] (14 practices)[c]	$791	$586	$204	-	-
Two years pre-IAH (14 practices)[c]	$849	$637	$212	$8 ($10)	1.0
Hospice					
Year 10 (1 practice)	$150	$103	$47	-$19 ($67)	-12.5
Year 9 (7 practices)	$183	$125	$58	-$11 ($24)	-6.7
Year 8 (7 practices)	$205	$131	$74	$5 ($18)	3.1
Year 7 (10 practices)	$169	$122	$47	-$23 ($14)	-14.1
Year 6 (12 practices)	$224	$130	$94	$30** ($14)	18.2
Year 5 (14 practices)	$167	$112	$55	-$1 ($11)	-0.8
Year 4 (14 practices)	$169	$113	$55	-$1 ($10)	-0.8
Year 3 (14 practices)	$162	$101	$61	$5 ($12)	3.0
Year 2 (14 practices)	$151	$84	$67	$10 ($11)	6.7
Year 1 (14 practices)	$162	$106	$56	$0 ($8)	-0.2
One year pre-IAH[b] (14 practices)[c]	$164	$108	$56	-	-
Two years pre-IAH (14 practices)[c]	$144	$100	$44	-$12 ($8)	-8.0
Outpatient					
Year 10 (1 practice)	$400	$500	-$100	-$39 ($74)	-37.0
Year 9 (7 practices)	$320	$527	-$207	-$134*** ($21)	-53.1
Year 8 (7 practices)	$374	$522	-$149	-$76*** ($16)	-30.1
Year 7 (10 practices)	$324	$463	-$139	-$68*** ($11)	-27.7
Year 6 (12 practices)	$361	$457	-$96	-$31 ($26)	-12.6
Year 5 (14 practices)	$319	$395	-$76	-$15 ($15)	-6.1
Year 4 (14 practices)	$296	$375	-$79	-$19 ($14)	-7.3
Year 3 (14 practices)	$294	$365	-$70	-$10 ($11)	-3.8
Year 2 (14 practices)	$285	$346	-$61	-$1 ($9)	-0.3
Year 1 (14 practices)	$263	$337	-$74	-$14* ($8)	-5.4
One year pre-IAH[b] (14 practices)[c]	$273	$334	-$61	-	-

Exhibit C.29 (*continued*)

Service type	IAH	Comparison	Difference (IAH - comparison)	Difference-in-differences estimated effect (SE)	Percentage effect[a]
Two years pre-IAH (14 practices)[c]	$268	$321	-$53	$8 ($8)	3.1
Clinician/supplier					
Year 10 (1 practice)	$1,116	$1,187	-$71	$18 ($90)	3.1
Year 9 (7 practices)	$948	$986	-$39	-$43 ($93)	-6.2
Year 8 (7 practices)	$908	$949	-$41	-$46 ($82)	-6.5
Year 7 (10 practices)	$798	$861	-$63	-$61 ($39)	-8.7
Year 6 (12 practices)	$803	$779	$24	$17 ($38)	2.6
Year 5 (14 practices)	$721	$724	-$2	-$57 ($35)	-8.0
Year 4 (14 practices)	$693	$675	$17	-$38 ($39)	-5.3
Year 3 (14 practices)	$723	$695	$28	-$27 ($32)	-3.8
Year 2 (14 practices)	$727	$692	$35	-$20 ($32)	-2.8
Year 1 (14 practices)	$744	$706	$38	-$18 ($24)	-2.5
One year pre-IAH[b] (14 practices)[c]	$772	$717	$55	-	-
Two years pre-IAH (14 practices)[c]	$775	$732	$42	-$13 ($10)	-1.8
Durable medical equipment					
Year 10 (1 practice)	$207	$135	$71	$4 ($47)	2.4
Year 9 (7 practices)	$164	$154	$10	-$22 ($18)	-15.7
Year 8 (7 practices)	$147	$142	$5	-$28** ($12)	-19.4
Year 7 (10 practices)	$146	$126	$20	-$13 ($16)	-9.0
Year 6 (12 practices)	$112	$101	$11	-$19** ($8)	-13.6
Year 5 (14 practices)	$89	$89	$0	-$38*** ($9)	-25.2
Year 4 (14 practices)	$87	$82	$5	-$33*** ($11)	-22.2
Year 3 (14 practices)	$100	$96	$4	-$35*** ($9)	-23.1
Year 2 (14 practices)	$110	$94	$16	-$22*** ($8)	-14.9
Year 1 (14 practices)	$145	$117	$28	-$10** ($5)	-6.7
One year pre-IAH[b] (14 practices)[c]	$163	$125	$38	-	-
Two years pre-IAH (14 practices)[c]	$170	$135	$35	-$3 ($5)	-2.1
Advanced imaging					
Year 10 (1 practice)	$400	$501	-$101	$22 ($56)	12.0
Year 9 (7 practices)	$320	$440	-$120	-$14 ($22)	-5.0
Year 8 (7 practices)	$318	$425	-$108	-$1 ($20)	-0.5
Year 7 (10 practices)	$299	$398	-$98	$9 ($14)	3.2

Exhibit C.29 (*continued*)

Service type	IAH	Comparison	Difference (IAH - comparison)	Difference-in-differences estimated effect (SE)	Percentage effect[a]
Year 6 (12 practices)	$313	$417	-$104	-$8 ($15)	-3.0
Year 5 (14 practices)	$278	$353	-$75	$1 ($12)	0.4
Year 4 (14 practices)	$246	$318	-$72	$4 ($14)	-3.4
Year 3 (14 practices)	$266	$348	-$82	-$6 ($14)	-2.9
Year 2 (14 practices)	$266	$351	-$85	-$8 ($15)	-2.1
Year 1 (14 practices)	$269	$355	-$86	-$10 ($13)	1.5
One year pre-IAH[b] (14 practices)[c]	$280	$356	-$76	-	-
Two years pre-IAH (14 practices)[c]	$297	$377	-$80	-$4 ($9)	-
Other imaging					
Year 10 (1 practice)	$653	$845	-$192	-$117 ($111)	-29.4
Year 9 (7 practices)	$468	$468	$0	-$141*** ($49)	-21.5
Year 8 (7 practices)	$522	$488	$34	-$107* ($60)	-16.3
Year 7 (10 practices)	$475	$475	$0	-$117** ($52)	-18.6
Year 6 (12 practices)	$482	$474	$8	-$88** ($39)	-15.1
Year 5 (14 practices)	$483	$450	$32	-$80 ($33)	-13.3
Year 4 (14 practices)	$478	$415	$63	-$50 ($41)	-8.3
Year 3 (14 practices)	$535	$476	$60	-$53 ($36)	-8.8
Year 2 (14 practices)	$576	$476	$100	-$13 ($30)	-2.2
Year 1 (14 practices)	$574	$455	$119	$6 ($20)	1.0
One year pre-IAH[b] (14 practices)[c]	$602	$490	$113	-	-
Two years pre-IAH (14 practices)[c]	$603	$484	$119	$6 ($14)	1.0
Testing					
Year 10 (1 practice)	$801	$935	-$134	-$41 ($100)	-7.1
Year 9 (7 practices)	$869	$684	$186	-$100 ($98)	-10.4
Year 8 (7 practices)	$888	$773	$115	-$171** ($80)	-17.6
Year 7 (10 practices)	$867	$705	$162	-$79 ($63)	-8.5
Year 6 (12 practices)	$1,010	$631	$379	$180 ($136)	21.0
Year 5 (14 practices)	$666	$562	$104	-$51 ($44)	-6.2
Year 4 (14 practices)	$636	$531	$105	-$49 ($45)	-6.1
Year 3 (14 practices)	$716	$593	$123	-$31 ($38)	-3.9
Year 2 (14 practices)	$796	$668	$128	-$26 ($29)	-3.2
Year 1 (14 practices)	$808	$684	$124	-$31 ($21)	-3.8

Exhibit C.29 (*continued*)

Service type	IAH	Comparison	Difference (IAH - comparison)	Difference-in-differences estimated effect (SE)	Percentage effect[a]
One year pre-IAH[b] (14 practices)[c]	$825	$671	$155	-	-
Two years pre-IAH (14 practices)[c]	$896	$680	$216	$61*** ($19)	7.5

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

Notes: This exhibit reports the regression-adjusted means of the IAH and matched comparison groups in each year. Estimates in each year use the sample of sites participating in that 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. We obtained regression-adjusted means by applying the estimated regression coefficients to the covariates of IAH beneficiaries in each demonstration year. Because of rounding, a difference-in-differences estimate displayed as zero might be shown alongside a percentage effect that exceeds zero. Estimated effects in Years 7, 8, and 9 are interpreted as effects of IAH during the first three years of the COVID-19 pandemic.

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

[a] We used the unadjusted IAH group mean in the year before the demonstration to calculate the percentage effect for each demonstration year. Exhibit C.22 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 performance year as the difference in means between IAH and comparison beneficiaries in that year minus the difference in the year before the demonstration.

[c] The pre-IAH spending amounts represent the spending of beneficiaries meeting IAH eligibility criteria for all 14 IAH practices in each pre-IAH year. The difference estimate for each pre-IAH year is not the same for IAH performance years with fewer than 14 practices.

SE = standard error.

Exhibit C.30. Estimated effect on primary and specialty care visits by performance year, Years 1 to 10

Service type	IAH	Comparison	Difference (IAH - comparison)	Difference-in-differences estimated effect (SE)	Percentage effect[a]
Primary care visits					
Year 10 (1 practice)	5,898	5,350	548	-3,217*** (561)	-42.4
Year 9 (7 practices)	8,937	5,582	3,356	-736 (653)	-8.0
Year 8 (7 practices)	9,205	5,407	3,798	-237 (407)	-2.6
Year 7 (10 practices)	10,669	5,255	5,413	1,466* (441)	16.1
Year 6 (12 practices)	11,789	5,622	6,166	1,982* (1,165)	21.0
Year 5 (14 practices)	10,742	5,408	5,334	855 (924)	8.9
Year 4 (14 practices)	10,710	5,293	5,418	938 (834)	9.8
Year 3 (14 practices)	10,269	5,269	5,000	521 (523)	5.4
Year 2 (14 practices)	10,124	5,275	4,849	369 (436)	3.9
Year 1 (14 practices)	9,637	5,190	4,446	-33 (270)	-0.3
One year pre-IAH[b] (14 practices)[c]	9,715	5,236	4,479	-	-
Two years pre-IAH (14 practices)[c]	9,460	5,310	4,150	-330 (161)**	-3.4
Specialty care visits					
Year 10 (1 practice)	6,733	9,703	-2,970	-645 (887)	-13.3
Year 9 (7 practices)	4,399	6,329	-1,930	-170 (368)	-4.1
Year 8 (7 practices)	4,648	6,227	-1,579	172 (389)	4.2
Year 7 (10 practices)	4,255	6,063	-1,808	-20 (212)	-0.5
Year 6 (12 practices)	4,134	6,114	-1,980	-443 (312)	-10.9
Year 5 (14 practices)	4,270	6,192	-1,922	-683* (384)	-15.7
Year 4 (14 practices)	4,219	5,974	-1,755	-516 (337)	-11.8
Year 3 (14 practices)	4,200	5,974	-1,774	-535* (293)	-12.3
Year 2 (14 practices)	4,365	6,096	-1,730	-491* (272)	-11.3
Year 1 (14 practices)	4,358	5,788	-1,431	-192 (285)	-4.4
One year pre-IAH[b] (14 practices)[c]	4,487	5,726	-1,239	-	-
Two years pre-IAH (14 practices)[c]	4,317	5,553	-1,236	3 (158)	0.1

Exhibit C.31. Estimated effect on hospital use by performance year, Years 1 to 10

Outcome	IAH	Comparison	Difference (IAH - comparison)	Difference-in-differences estimated effect (SE)	Percentage effect[a]
Number of hospital admissions per 1,000 beneficiaries per year					
Year 10 (1 practice)	2,483	1,771	712	917*** (248)	65.0
Year 9 (7 practices)	1,927	2,186	-259	127 (84)	7.1
Year 8 (7 practices)	1,891	2,187	-296	100 (90)	5.6
Year 7 (10 practices)	1,639	2,025	-386	-33 (81)	-1.9
Year 6 (12 practices)	1,761	2,076	-315	25 (66)	1.4
Year 5 (14 practices)	1,785	2,118	-332	-129** (62)	-7.2
Year 4 (14 practices)	1,630	1,935	-304	-101 (71)	-5.7
Year 3 (14 practices)	1,761	2,049	-287	-84 (64)	-4.7
Year 2 (14 practices)	1,786	2,020	-234	-30 (52)	-1.7
Year 1 (14 practices)	1,828	2,081	-252	-49 (36)	-2.8
One year pre-IAH[b] (14 practices)[c]	1,898	2,102	-203	-	-
Two years pre-IAH (14 practices)[c]	1,979	2,216	-237	-33 (29)	-1.9
Number of hospital admissions preceded by an ED visit per 1,000 beneficiaries per year					
Year 10 (1 practice)	2,185	1,552	633	782*** (228)	63.3
Year 9 (7 practices)	1,629	1,777	-149	22 (81)	1.5
Year 8 (7 practices)	1,571	1,763	-192	-16 (88)	-1.1
Year 7 (10 practices)	1,372	1,634	-262	-107 (77)	-7.4
Year 6 (12 practices)	1,421	1,598	-178	-33 (61)	-2.3
Year 5 (14 practices)	1,452	1,636	-184	-149** (62)	-10.4
Year 4 (14 practices)	1,298	1,473	-175	-141** (61)	-9.8
Year 3 (14 practices)	1,410	1,572	-162	-127** (49)	-8.8
Year 2 (14 practices)	1,427	1,527	-99	-65 (45)	-4.5
Year 1 (14 practices)	1,499	1,603	-104	-69** (31)	-4.8
One year pre-IAH[b] (14 practices)[c]	1,573	1,608	-34	-	-
Two years pre-IAH (14 practices)[c]	1,661	1,711	-49	-14 (26)	-1.0
Number of potentially avoidable hospital admissions per 1,000 beneficiaries per year					
Year 10 (1 practice)	402	308	94	192* (105)	52.3
Year 9 (7 practices)	348	399	-51	17 (34)	3.7
Year 8 (7 practices)	365	396	-31	39 (44)	8.7
Year 7 (10 practices)	306	395	-89	-27 (31)	-5.9
Year 6 (12 practices)	385	454	-69	-8 (27)	-1.9
Year 5 (14 practices)	457	544	-87	-55** (24)	-12.1
Year 4 (14 practices)	361	452	-90	-58** (23)	-12.8

Exhibit C.31 (*continued*)

Outcome	IAH	Comparison	Difference (IAH - comparison)	Difference-in-differences estimated effect (SE)	Percentage effect[a]
Year 3 (14 practices)	412	481	-68	-36* (18)	-7.9
Year 2 (14 practices)	433	470	-37	-5 (19)	-1.1
Year 1 (14 practices)	463	508	-45	-13 (12)	-2.8
One year pre-IAH[b] (14 practices)[c]	487	519	-32	-	-
Two years pre-IAH (14 practices)[c]	510	569	-58	-26* (14)	-5.8
Probability of having a qualifying hospital discharge and unplanned readmission within 30 days of discharge					
Year 10 (1 practice)	23.35	12.85	10.51	9.80** (4.54)	56.9
Year 9 (7 practices)	15.69	16.95	-1.27	1.42* (0.83)	7.2
Year 8 (7 practices)	15.62	17.33	-1.71	0.76 (1.31)	3.9
Year 7 (10 practices)	15.74	18.60	-2.86	-0.55 (0.99)	-2.8
Year 6 (12 practices)	15.91	18.17	-2.26	-0.02 (1.05)	-0.1
Year 5 (14 practices)	15.86	18.14	-2.29	-1.47 (1.05)	-9.2
Year 4 (14 practices)	13.65	15.97	-2.32	-1.50 (0.98)	-9.4
Year 3 (14 practices)	16.18	18.70	-2.52	-1.70** (0.77)	-10.6
Year 2 (14 practices)	16.32	18.14	-1.82	-1.00 (0.98)	-6.3
Year 1 (14 practices)	17.93	18.63	-0.70	0.12 (0.66)	0.7
One year pre-IAH[b] (14 practices)[c]	17.78	18.60	-0.82	-	-
Two years pre-IAH (14 practices)[c]	19.88	20.89	-1.02	-0.20 (0.66)	-1.3

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

Notes: This exhibit reports the regression-adjusted means of the IAH and matched comparison groups in each year. Estimates in each year use the sample of sites participating in that 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. We obtained regression-adjusted means by applying the estimated regression coefficients to the covariates of IAH beneficiaries in each demonstration year. Because of rounding, a difference-in-differences estimate displayed as zero might be shown alongside a percentage effect that exceeds zero. Estimated effects in Years 7, 8, and 9 are interpreted as effects of IAH during the first three years of the COVID-19 pandemic.

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

[a] We used the unadjusted IAH group mean in the year before the demonstration to calculate the percentage effect for each demonstration year. Exhibit C.22 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 performance year as the difference in means between IAH and comparison beneficiaries in that year minus the difference in the year before the demonstration.

[c] The pre-IAH spending amounts represent the spending of beneficiaries meeting IAH eligibility criteria for all 14 IAH practices in each pre-IAH year. The difference estimate for each pre-IAH year is not the same for IAH performance years with fewer than 14 practices.

SE = standard error.

Exhibit C.32. Estimated effect on outpatient ED visits and potentially avoidable outpatient ED visits by performance year, Years 1 to 10

Outcome	IAH	Comparison	Difference (IAH - comparison)	Difference-in-differences estimated effect (SE)	Percentage effect[a]
Number of outpatient ED visits per 1,000 beneficiaries per year					
Year 10 (1 practice)	1,269	924	345	385** (191)	57.6
Year 9 (7 practices)	1,555	1,680	-125	-66 (99)	-4.5
Year 8 (7 practices)	1,611	1,675	-64	-4 (74)	-0.3
Year 7 (10 practices)	1,332	1,450	-118	-54 (66)	-3.8
Year 6 (12 practices)	1,755	1,797	-42	-25 (102)	-1.8
Year 5 (14 practices)	1,683	1,732	-48	-55 (99)	-3.8
Year 4 (14 practices)	1,613	1,662	-49	-56 (80)	-3.9
Year 3 (14 practices)	1,701	1,711	-9	-17 (72)	-1.2
Year 2 (14 practices)	1,639	1,561	78	70 (54)	4.9
Year 1 (14 practices)	1,461	1,488	-27	-34 (53)	-2.4
One year pre-IAH[b] (14 practices)[c]	1,503	1,496	7	-	-
Two years pre-IAH (14 practices)[c]	1,475	1,470	4	-2 (37)	-0.2
Number of potentially avoidable outpatient ED visits per 1,000 beneficiaries per year					
Year 10 (1 practice)	227	139	88	6 (93)	4.9
Year 9 (7 practices)	209	213	-4	13 (19)	6.9
Year 8 (7 practices)	223	212	11	27 (17)	14.2
Year 7 (10 practices)	180	196	-16	3 (15)	1.4
Year 6 (12 practices)	254	254	0	17 (17)	9.0
Year 5 (14 practices)	245	263	-18	-7 (13)	-3.6
Year 4 (14 practices)	214	243	-29	-18 (13)	-9.3
Year 3 (14 practices)	223	229	-6	5 (14)	2.6
Year 2 (14 practices)	223	216	6	18 (11)	9.5
Year 1 (14 practices)	198	206	-7	4 (11)	1.9
One year pre-IAH[b] (14 practices)[c]	202	213	-11	-	-
Two years pre-IAH (14 practices)[c]	203	208	-4	6 (13)	3.5

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

Notes: This exhibit reports the regression-adjusted means of the IAH and matched comparison groups in each year. Estimates in each year use the sample of sites participating in that 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. We obtained regression-adjusted means by applying the estimated regression

Exhibit C.32 (*continued*)

coefficients to the covariates of IAH beneficiaries in the latest demonstration year. Because of rounding, a difference-in-differences estimate displayed as zero might be shown alongside a percentage effect that exceeds zero. Estimated effects in Years 7, 8, and 9 are interpreted as effects of IAH during the first three years of the COVID-19 pandemic.

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

[a] We used the unadjusted IAH group mean in the year before the demonstration to calculate the percentage effect for each demonstration year. Exhibit C.22 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 performance year as the difference in means between IAH and comparison beneficiaries in that year minus the difference in the year before the demonstration.

[c] The pre-IAH spending amounts represent the spending of beneficiaries meeting IAH eligibility criteria for all 14 IAH practices in each pre-IAH year. The difference estimate for each pre-IAH year is not the same for IAH performance years with fewer than 14 practices.

SE = standard error.

Exhibit C.33. Estimated effect on the probability of entering a nursing home by performance year, Years 1 to 10

	IAH	Comparison	Difference (IAH - comparison)	Difference-in-differences estimated effect (SE)	Percentage effect[a]
Year 10 (1 practice)	2.99	14.79	-11.79	-7.33** (3.41)	-88.2
Year 9 (7 practices)	6.07	13.80	-7.73	0.12 (1.05)	1.4
Year 8 (7 practices)	6.10	12.38	-6.28	1.59* (0.82)	18.5
Year 7 (10 practices)	5.50	13.13	-7.63***	-0.07 (0.62)	-0.8
Year 6 (12 practices)	6.58	12.02	-5.44	2.38*** (0.60)	34.4
Year 5 (14 practices)	6.70	12.89	-6.20	0.87 (0.62)	12.4
Year 4 (14 practices)	6.51	12.78	-6.26	0.15 (0.51)	2.1
Year 3 (14 practices)	7.44	13.26	-5.82	0.59 (0.49)	8.4
Year 2 (14 practices)	7.10	12.76	-5.66	0.75 (0.47)	10.7
Year 1 (14 practices)	7.46	13.49	-6.03	0.38 (0.47)	5.5
One year pre-IAH[b] (14 practices)[c]	7.88	14.29	-6.41	-	-
Two years pre-IAH (14 practices)[c]	8.03	15.14	-7.11	-0.70 (0.49)	-10.1

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

Notes: This exhibit reports the regression-adjusted means of the IAH and matched comparison groups in each year. Estimates in each year use the sample of sites participating in that 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. We obtained regression-adjusted means by applying the estimated regression coefficients to the covariates of IAH beneficiaries in each demonstration year. Because of rounding, a difference-in-differences estimate displayed as zero might be shown alongside a percentage effect that exceeds zero. Estimated effects in Years 7, 8, and 9 are interpreted as effects of IAH during the first three years of the COVID-19 pandemic.

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

[a] We used the unadjusted IAH group mean in the year before the demonstration to calculate the percentage effect for each demonstration year. Exhibit C.22 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 performance year as the difference in means between IAH and comparison beneficiaries in that year minus the difference in the year before the demonstration.

[c] The pre-IAH spending amounts represent the spending of beneficiaries meeting IAH eligibility criteria for all 14 IAH practices in each pre-IAH year. The difference estimate for each pre-IAH year is not the same for IAH performance years with fewer than 14 practices.

SE = standard error.

Exhibit C.34. Estimated effect on the probability of dying by performance year, Years 1 to 10

	IAH	Comparison	Difference (IAH - comparison)	Difference-in-differences estimated effect (SE)	Percentage effect[a]
Year 10 (1 practice)	16.12	26.61	-10.49	-1.40 (2.49)	-6.7
Year 9 (7 practices)	14.10	23.31	-9.21	-2.43*** (0.83)	-15.0
Year 8 (7 practices)	15.69	24.75	-9.06	-2.26*** (0.75)	-14.0
Year 7 (10 practices)	15.83	24.92	-9.10	-2.42*** (0.88)	-14.4
Year 6 (12 practices)	15.17	21.11	-5.94	0.56 (0.86)	3.8
Year 5 (14 practices)	13.90	20.58	-6.68	-1.20* (0.70)	-8.1
Year 4 (14 practices)	13.74	20.52	-6.77	-1.29** (0.63)	-8.7
Year 3 (14 practices)	15.43	19.88	-4.44	1.04 (0.68)	7.0
Year 2 (14 practices)	14.71	19.86	-5.15	0.33 (0.67)	2.2
Year 1 (14 practices)	15.36	20.83	-5.47	0.02 (0.56)	0.1
One year pre-IAH[b] (14 practices)[c]	15.54	21.03	-5.48	-	-
Two years pre-IAH (14 practices)[c]	14.94	22.05	-7.11	-1.62*** (0.64)	-10.9

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

Notes: This exhibit reports the regression-adjusted means of the IAH and matched comparison groups in each year. Estimates in each year use the sample of sites participating in that 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. We obtained regression-adjusted means by applying the estimated regression coefficients to the covariates of IAH beneficiaries in each demonstration year. Because of rounding, a difference-in-differences estimate displayed as zero might be shown alongside a percentage effect that exceeds zero. Estimated effects in Years 7, 8, and 9 are interpreted as effects of IAH during the first three years of the COVID-19 pandemic.

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

[a] We used the unadjusted IAH group mean in the year before the demonstration to calculate the percentage effect for each demonstration year. Exhibit C.22 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 performance year as the difference in means between IAH and comparison beneficiaries in that year minus the difference in the year before the demonstration.

[c] The pre-IAH spending amounts represent the spending of beneficiaries meeting IAH eligibility criteria for all 14 IAH practices in each pre-IAH year. The difference estimate for each pre-IAH year is not the same for IAH performance years with fewer than 14 practices.

SE = standard error.

Exhibit C.35. Estimated effect on days at home by performance year, Years 1 to 10

	IAH	Comparison	Difference (IAH - comparison)	Difference-in-differences estimated effect (SE)	Percentage effect[a]
Year 10 (1 practice)	88.58	85.07	3.51	-1.09 (1.48)	-1.2
Year 9 (7 practices)	90.65	86.62	4.03	0.27 (0.60)	0.3
Year 8 (7 practices)	90.74	86.99	3.75	0.00 (0.52)	0.0
Year 7 (10 practices)	91.54	87.33	4.21	0.53 (0.43)	0.6
Year 6 (12 practices)	90.61	87.58	3.03	0.53 (.43)	-0.8
Year 5 (14 practices)	90.18	86.85	3.33	-0.64* (.34)	-0.4
Year 4 (14 practices)	90.50	87.03	3.47	0.30% (0.33)	-0.4
Year 3 (14 practices)	89.43	86.25	3.18	0.44% (0.34)	-0.6
Year 2 (14 practices)	89.37	86.37	2.99	0.15% (0.26)	-0.9
Year 1 (14 practices)	89.38	86.26	3.11	-0.04% (0.29)	-0.6
One year pre-IAH[b] (14 practices)[c]	89.28	86.25	3.03	-	-
Two years pre-IAH (14 practices)[c]	88.64	85.45	3.19	0.16 (0.24)	0.2

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

Notes: This exhibit reports the regression-adjusted means of the IAH and matched comparison groups in each year. Estimates in each year use the sample of sites participating in that 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. We obtained regression-adjusted means by applying the estimated regression coefficients to the covariates of IAH beneficiaries in each demonstration year. Because of rounding, a difference-in-differences estimate displayed as zero might be shown alongside a percentage effect that exceeds zero. Estimated effects in Years 7, 8, and 9 are interpreted as effects of IAH during the first three years of the COVID-19 pandemic.

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

[a] We used the unadjusted IAH group mean in the year before the demonstration to calculate the percentage effect for each demonstration year. Exhibit C.22 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 performance year as the difference in means between IAH and comparison beneficiaries in that year minus the difference in the year before the demonstration.

[c] The pre-IAH spending amounts represent the spending of beneficiaries meeting IAH eligibility criteria for all 14 IAH practices in each pre-IAH year. The difference estimate for each pre-IAH year is not the same for IAH performance years with fewer than 14 practices.

SE = standard error.

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). "Risk Adjustment." Baltimore, MD: CMS, 2019. Available at https://www.cms.gov/Medicare/Health-Plans/MedicareAdvtgSpecRateStats/Risk-Adjustors.html. Accessed June 6, 2023.

Centers for Medicare & Medicaid Services (CMS). "Independence at Home Demonstration Revised Actuarial Shared Savings Methodology." Baltimore, MD: CMS, 2021. Available at https://innovation.cms.gov/media/document/iah-rev-act-meth-specs. Accessed June 4, 2025.

Centers for Medicare & Medicaid Services (CMS). "Shared Savings Program Fast Facts – As of January 1, 2023." Baltimore, MD: CMS, 2023a. Available at https://www.cms.gov/files/document/2023-shared-savings-program-fast-facts.pdf. Accessed June 4, 2025.

Centers for Medicare & Medicaid Services (CMS). "Independence at Home Demonstration Performance Year 8 Results." Baltimore, MD: CMS, 2023b. Available at https://innovation.cms.gov/media/document/iah-py8-fs. Accessed June 4, 2025.

Centers for Medicare & Medicaid Services (CMS). "QPP 2021 Participation Results At-A-Glance." Baltimore, MD: CMS, 2023c. Available at: https://qpp.cms.gov/resources/document/687503c5-dfd8-46c8-80f6-7e979b9e7d1c. Accessed June 6, 2025.

Cottrill, A., J. Cubanski, and T. Neuman. "What to Know About How Medicare Pays Physicians." KFF, March 24, 2025. Available at https://www.kff.org/medicare/issue-brief/what-to-know-about-how-medicare-pays-physicians/#. Accessed June 4, 2025.

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.

Hannan, E., Y. Yu, K. Cozzens, and B. Anderson. "The Neighborhood Atlas Area Deprivation Index For Measuring Socioeconomic Status: An Overemphasis On Home Value." *Health Affairs*, vol. 42, no. 5, 2023, pp. 702–709.

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. Accessed July 2, 2024.

Office of the Inspector General. "Consumer Alert: Urinary Catheter Scams." 2024. https://oig.hhs.gov/fraud/consumer-alerts/consumer-alert-catheter-scam/.

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.

Schochet, P.Z. "Statistical power for random assignment evaluations of education programs." *Journal of Educational and Behavioral Statistics*, vol. 33, no. 1, 2008, pp. 62-87.

Stuart, E.A. "Matching Methods for Causal Inference: A Review and Look Forward." *Statistical Science*, vol. 21, no. 1, 2010, pp. 1–25.

Appendix D.

Technical Appendix for the Three-Year Analysis of IAH Beneficiaries New to HBPC

1. Overview of the three-year analysis of IAH beneficiaries new to HBPC

As we described briefly in Chapter 1 of Appendix C, the three-year analysis examined the effects of HBPC provided by IAH practices on spending, hospital use, and health outcomes by studying beneficiaries who were new to HBPC relative to a comparison group of similar beneficiaries who did not receive HBPC for at least another year, if at all. We conducted this analysis using beneficiaries who started receiving HBPC from IAH practices when the IAH payment incentive was in effect. Therefore, it reflects the combined effect of both components of the IAH demonstration, as shown in Section 1.1 of the report.

Exhibit D.1 describes what the difference-in-differences estimate from the three-year analysis measures. The important difference between IAH and comparison beneficiaries is the choice to begin receiving HBPC from an IAH practice during a period when the IAH practice could have earned incentive payments. Thus, the difference-in-differences analysis we describe in Chapter 4 of this appendix estimates the effect of newly receiving HBPC from an IAH practice that had the opportunity to earn IAH incentive payments.

Exhibit D.1. Results for outcomes estimated by the three-year analysis of IAH beneficiaries new to HBPC

Intervention received before starting HBPC (pre-period)			Intervention received after starting HBPC (post-period)			Intervention tested by the difference-in-differences analysis (6) – (3) (7)
IAH beneficiaries (1)	Comparison beneficiaries (2)	IAH-Comparison difference (1) – (2) (3)	IAH beneficiaries (4)	Comparison beneficiaries (5)	IAH-Comparison difference (4) – (5) (6)	
Did *not* receive HBPC, and the beneficiary's provider could *not* have received IAH incentive payments		None	Received HBPC from an IAH practice during a time when the practice could have received IAH incentive payments	Did *not* receive HBPC, and the beneficiary's provider could *not* have received IAH incentive payments	IAH beneficiaries receive HBPC from IAH practices eligible for incentive payments, while comparisons do not	The DID estimate reflects the effect of newly receiving HBPC from an IAH practice during a time when the practice could have received IAH incentive payments

HBPC = home-based primary care.

We used a matched retrospective cohort design with an intent-to-treat (ITT) analysis, estimating effects using difference-in-differences regression. The sections that follow describe the sample construction and effect estimation processes used in the three-year analysis in more detail.

2. Methods for identifying the IAH beneficiaries

The intervention group for the three-year analysis comprised beneficiaries who: (1) began receiving HBPC from IAH practices for the first time during the IAH demonstration and (2) met IAH eligibility requirements at the time of their first HBPC visit. IAH eligibility largely followed the criteria used in the performance-year analysis, with some additional restrictions.

2.1. Identifying IAH beneficiaries for the intervention group

For this analysis, we identified a pooled intervention group of beneficiaries who met our eligibility criteria during each year from 2013 to 2021, which differed from the performance-year analysis, where we identified an intervention group separately in each demonstration year. There were nine panel years that corresponded to calendar years 2013 to 2021 (Exhibit D.2).

In each of the nine years, a new group of IAH beneficiaries entered the sample. The date of the beneficiary's first HBPC visit from an IAH practice was the *index date*. The index date was the first day of the beneficiary's three-year follow-up period. In this analysis, we refer to the three years before a beneficiary's index date as the pre-period and the three years after the index date as the follow-up period.

New IAH beneficiaries were new to both HBPC and to the IAH practice based on a two-year lookback, defined as having had no visits with an IAH clinician and no home visits from any primary care clinician or specialist physician within two years before the index date (including home visits from rural health clinics or federally qualified health centers). We used a two-year look-back rather than (1) one year, which could have included some beneficiaries with previous HBPC experience, or (2) a longer period of time, which would have limited sample sizes.

Exhibit D.2. Mapping panel years to IAH demonstration years and IAH participation to determine which states to include in each panel year

Panel year	Calendar year	Overlap with IAH demonstration	States included in the panel year for identifying IAH beneficiaries, based on location of practices that participated in IAH in the following demonstration year
1	2013	IAH Years 1 and 2	DE, DC, FL, MA, MI, NC, NY, OH, OR, PA, TX, VA, WI
2	2014	IAH Years 2 and 3	DE, DC, FL, MA, MI, NC, NY, OH, OR, PA, TX, VA, WI
3	2015	IAH Years 3 and 4 plus gap between 3 and 4	DE, DC, FL, MA, MI, NC, NY, OH, OR, PA, TX, VA, WI
4	2016	IAH Years 4 and 5	DE, DC, FL, MA, MD, MI, NC, NY, OH, OR, PA, TX, VA, WI
5	2017	IAH Year 5	DE, DC, FL, MA, MD, MI, NC, NY, OR, PA, TX, VA, WI
6	2018	Gap between IAH Years 5 and 6*	DE, DC, FL, MA, MD, MI, NC, NY, OR, PA, TX, VA, WI
7	2019	IAH Year 6	DE, DC, FL, MD, MI, NY, OR, PA, TX, VA, WI
8	2020	IAH Year 7	FL, MD, MI, NY, PA, TX, VA, WI
9	2021	IAH Year 8	FL, MD, MI, NY, PA, TX, VA, WI

All beneficiaries had to meet the following IAH eligibility criteria to be eligible for selection as an intervention beneficiary in a given panel year. This approach generally mirrored the approach that we used in the performance-year analysis (see Appendix C for details):

- Had two or more chronic conditions as of the index date, as measured by chronic condition indicators in the Chronic Conditions Warehouse
- Had two or more activities of daily living (ADLs) requiring human assistance in the first 12 months after the index date according to the Outcome and Assessment Information Set (OASIS), Minimum Data Set (MDS), or Inpatient Rehabilitation Facility Patient Assessment Instrument (IRF-PAI)[43]
- Had an inpatient hospital admission or observation stay within 12 months prior to the index date[44]
- Used acute or subacute rehabilitation services within 12 months prior to the index date (home health, skilled nursing facility [SNF], or inpatient rehabilitation hospital)
- Were not in hospice within 30 days of the index date.
- Were IAH-eligible as of the index date, which was the first HBPC visit from an IAH practice. In the performance-year analysis, we identified the first month that a beneficiary was IAH-eligible during a given performance year.

In addition to meeting IAH eligibility criteria, all beneficiaries had to meet the following criteria to be eligible for selection as an intervention beneficiary in a given panel year:

- Lived in an IAH state at the beginning of a panel year based on their state information at the end of the previous year (Exhibit D.2)
- Had continuous Medicare fee-for-service (FFS) status that began three years before the index date
- Had continuous Medicare FFS status in the panel year (unless the beneficiary died in that year)
- Were eligible to receive HBPC based on residing in the community as of the index date and the day after the index date according to the Timeline file

For intervention beneficiaries who remained alive and enrolled in FFS Medicare for at least another three months after the index date, a beneficiary had to receive a second visit from the IAH practice during the first follow-up year. This approach mimics how the intervention group was identified for the performance-year analysis.

The last year that an IAH beneficiary could enter the sample was the calendar year before their IAH practice left the demonstration. For example, three practices left the demonstration after Year 7, which ended on December 31, 2020. We identified beneficiaries for the intervention group for these three practices based on index dates as late as December 31, 2019. This approach ensured that all beneficiaries in the sample had at least one year during the follow-up period when they were exposed to both components of the intervention—HBPC from the IAH practice and the IAH payment incentive.

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

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

The last calendar year when an IAH beneficiary could enter the sample was 2021 to allow a full year of observation for data available through 2022 at the time of the analysis.

We allowed a beneficiary to enter the intervention group only once. For example, if an IAH-eligible beneficiary was new to HBPC at the IAH practice in February 2013, received HBPC through December 2013, and then had no more HBPC for more than three years before receiving home visits from clinicians with the practice in July 2017, this person would enter the sample in February 2013 only—not also July 2017. However, a beneficiary who received HBPC from an IAH practice but did not qualify for the intervention group based on other eligibility criteria could have later entered the intervention group after two years of not receiving any HBPC.

We used an intent-to-treat design. IAH beneficiaries remained in the intervention group until they died, left FFS Medicare, or their three-year follow-up period ended—even if they stopped receiving HBPC from an IAH practice. We analyzed retention to see how often IAH beneficiaries remained alive and in FFS but stopped receiving HBPC from an IAH practice after the first year, when 100 percent of IAH beneficiaries had an HBPC visit from an IAH practice. Among those who were alive and in FFS at the beginning of the second year, 63 percent had an HBPC visit from an IAH practice. The figure was 52 percent among those who were alive and in FFS at the beginning of the third year.

2.2. Implications of the IAH demonstration timeline for sample entry and the follow-up period

The timing of the demonstration, particularly the gaps in the demonstration and practices' staggered exits, affected when we allowed beneficiaries to enter the intervention group, as well as when we measured outcomes for beneficiaries in the study sample.

We treated 2013 as the first year of the study period and 2022 as the last year of the study period. We did not include 2012 because the demonstration began in mid-2012. Claims data for 2023 were not available when we constructed the sample and outcome data for the three-year analysis. This meant that 2021 was the last year an IAH beneficiary could enter the intervention group. This decision rule resulted in some beneficiaries having less than three years of follow-up time—beneficiaries who joined the sample in 2020 had only two full years of follow-up and beneficiaries who joined the sample in 2021 had only one full year. However, allowing beneficiaries to have one or two years of follow-up meant that we could retain a larger sample than if we required all beneficiaries to have three years of follow-up. Requiring three years of follow-up would have meant allowing no more beneficiaries to join the sample after December 2019, rather than December 2021, which would exclude beneficiaries who joined in four of the 10 demonstration years.

In keeping with our ITT design, we followed all intervention beneficiaries up to three years after their index date, regardless of whether the beneficiary continued to receive HBPC or the beneficiary's practice remained in the demonstration. Likewise, comparison beneficiaries who began to receive HBPC from a non-IAH provider during their follow-up period remained in the sample. Beneficiaries who died or left FFS Medicare over the course of a follow-up year had their continuous outcomes (such as spending and hospital use) annualized (spread over a full 12 months) and then were downweighted in the regressions using observability weights. For example, if a beneficiary died or left FFS Medicare 14 months after the

index date, then we defined the outcome in the second follow-up year based on two months of data, annualized the result to reflect 12 months, and weighted that outcome accordingly.

The demonstration paused between Years 3 and 4 and Years 5 and 6; we call the periods when the demonstration was paused gaps. The first gap lasted from June to September 2015 for all practices except the Richmond-based consortium, which had a gap in September 2015 only. The gap between Years 5 and 6 lasted from October 2017 to December 2018 for all practices. During gaps in the IAH payment incentive, we allowed beneficiaries to join the sample and continued to measure outcomes for beneficiaries already in the sample. During interviews with IAH practices, most practices reported no major changes in care processes during the gap in the payment incentive, and IAH practices provided HBPC during this time.

Throughout the demonstration, practices exited at different times. Larger practices and practices that remained in the demonstration longer had more influence on the results because they accounted for more of the sample. We continued to measure outcomes for IAH beneficiaries after their practice left the demonstration—that is, after their practice was no longer eligible to receive an incentive payment for providing HBPC as part of participating in IAH—for the following reasons:

- IAH practices reported that their basic model of HBPC was unchanged during the demonstration, and therefore, we expected that it continued after they exited the demonstration.
- During interviews in 2019, most practices reported that they had applied changes they made because of IAH to deliver better care to all their patients, not just IAH beneficiaries, which lessened the chance practices would change those approaches simply because they left the IAH demonstration.

In sum, it is unlikely that measuring outcomes for IAH beneficiaries after their practice left the demonstration diluted the estimated effects.

3. Methods for identifying the comparison group

3.1. Potential comparison group

To identify the pool of potential comparison group beneficiaries, we used eligibility criteria similar to the intervention group, including:

- Met the same requirements as the intervention group for IAH eligibility, including no HBPC use or visits to an IAH practice in the past two years
- Lived in the catchment area of an IAH practice, which we defined using IAH beneficiaries' ZIP codes
- Had an ambulatory visit with a primary care or specialty care clinician (used to identify an index date)
- Did not receive HBPC from any provider for 12 months after entering the sample and did not visit an IAH practice during this 12-month period
- Did not meet eligibility criteria for the IAH group and start HBPC at an IAH practice at any point in the study period before or after they were considered for inclusion in the potential comparison group

Unlike the IAH beneficiaries, potential comparison beneficiaries did not receive HBPC, so we could not use the first HBPC visit to set an index date to determine the start of the follow-up period. We used separate criteria to set index dates for potential comparison beneficiaries. Given that generally sicker beneficiaries enter HBPC, leading to more hospital admissions and higher spending, it was essential that matched comparison beneficiaries be at a point in their health trajectory similar to the IAH beneficiaries when they began HBPC. If we chose a single index date at random for each potential comparison for the purpose of defining the baseline period, we could miss the point at which that beneficiary was an ideal match for an IAH beneficiary. Therefore, we created multiple index dates for beneficiaries in the potential comparison group: each month in which a beneficiary met the eligibility criteria and had an ambulatory visit to a primary care or specialty care clinician (physician, nurse practitioner [NP], or physician assistant [PA]) represented a different version of that beneficiary in the potential comparison pool. The index date was defined as the potential comparison beneficiary's first ambulatory visit in the month, mirroring the definition of the index date for IAH beneficiaries. Allowing multiple index dates was not the only way we tried to select comparison beneficiaries based on their health trajectory. As we describe in the next section, we included many variables in propensity score matching to address this issue, such as number of months since the most recent hospital admission.

A potential comparison beneficiary could have up to 12 versions per year, each with a different index date used to better match with an IAH beneficiary. Also, potential comparison beneficiaries could have versions in different years in which they are eligible (2013 to 2021). As part of our comparison group selection approach, we selected at most one index date for each comparison beneficiary, and a comparison beneficiary could be matched to no more than one IAH beneficiary. The version of the index date we selected for the matched comparison group (if any) was the one that best matched to an IAH beneficiary on key characteristics, such as recent health trends and proximity of the index date to the IAH beneficiary's index date.

In keeping with the intent-to-treat design and our approach for the intervention group, beneficiaries selected as matched comparisons remained in the comparison group until they died, left FFS, or their follow-up period ended. We analyzed retention to see how often the matched comparison group started receiving HBPC from a non-IAH clinician after the first follow-up year. By design, none of the comparison beneficiaries had an HBPC visit in the first follow-up year, and none of them had an HBPC visit from an IAH clinician in any follow-up year. Among those who were alive and in FFS at the beginning of the second year, 11 percent had an HBPC visit from a non-IAH clinician. The figure was 15 percent among those who were alive and in FFS at the beginning of the third year. It is possible that the estimated effects in the second and third years were slightly underestimated because of the comparison beneficiaries who began HBPC during those years. However, it is unlikely that this HBPC use substantially affected results for the three-year analysis for two reasons. First, HBPC users were a small minority of the full comparison group, most of whom never received HBPC over the three-year follow-up period. Second, these HBPC users did not visit IAH practices. Some or perhaps many of them had HBPC from a clinician who does not specialize in HBPC, like IAH practices do. As a result, the type and perhaps the quality of HBPC they received likely differed from the care the IAH beneficiaries received. As we described in Chapter 12 of the report, HBPC visits are generally provided by clinicians who do not specialize in HBPC. For a discussion about analyses testing sensitivity to how we applied the intent-to-treat design to the intervention group, see Section 4.2.7 of this appendix (Exhibits D.6 and D.7).

3.2. Propensity score matching methods

As in the performance-year analysis, we used propensity score matching to create a comparison group of beneficiaries who did not begin HBPC but were similar in observable characteristics to IAH beneficiaries who started HBPC during the demonstration. Matching had the same goals in the three-year analysis as in the performance-year analysis: to reduce the effect of selection into HBPC on our estimates by comparing beneficiaries who started HBPC at IAH practices during the demonstration to otherwise similar beneficiaries who did not and to improve the statistical efficiency of our estimates by selecting the largest comparison group that met our criteria.

We selected matched comparison beneficiaries for each IAH beneficiary based on their similarity in the year before the IAH beneficiary started HBPC. Because comparison beneficiaries never started HBPC in the first year, we considered several different index dates for each potential comparison beneficiary, as described above. Each index date defines a snapshot of a potential comparison beneficiary based on health and demographic characteristics at that point in time. Over time, beneficiaries' health and medical complexity change, so potential comparisons could be more similar to the IAH beneficiaries at some points in time than others. To determine which snapshot, if any, was the best match for an IAH beneficiary, we used an approach called GroupMatch (Pimentel et al. 2020) that is designed for this type of problem. GroupMatch selects at most one snapshot of each potential comparison beneficiary to include in the comparison group, based on the beneficiaries' similarity on matching characteristics.

Unlike in the performance-year analysis, we did not match separately by IAH practice and by year; we estimated a single propensity score model across beneficiaries from all practices in all years. However, we applied exact-match restrictions that ensured that IAH beneficiaries were matched to comparison beneficiaries living in their practice's service area. Also, because beneficiaries in the three-year analysis

started HBPC between 2013 and 2021, we required comparison beneficiaries to have an index date within three months of the IAH beneficiary's HBPC index date.

We used the same measures of health status and demographic characteristics in matching for the three-year analysis as we did in the performance-year analysis. However, we added some variables. In one case, the variable was not available at the time we conducted matching for the performance-year analysis in earlier years of the evaluation. In the other cases, we used these variables in the three-year analysis but not the performance-year analysis because of the difference in the two study designs. Adding the following variables to matching in the three-year analysis helped us to improve balance on health trajectories, which was essential because we followed individual beneficiaries over time in this analysis:

- We matched beneficiaries exactly on whether they died within three months of starting HBPC (or within three months of their index date, for comparison beneficiaries).
- We matched on a claims-based frailty index. We did not use this index to select the comparison group for the performance-year analysis because the index had not yet been developed when we established the matching approach for the performance-year analysis early in the demonstration. The claims-based frailty index (CFI) is a continuous frailty score that ranges from 0 to 1. It was developed by Kim et al. (2018) using the following Medicare FFS claims: inpatient, outpatient, SNF, home health agency, carrier, and durable medical equipment. The CFI is based on 93 indicators created using ICD-9-CM/ICD-10-CM diagnosis codes (52 indicators) and CPT/HCPCS codes (41 indicators) in the prior 12 months. Examples of these indicators include use of durable medical equipment, home health services, and mobility devices along with frailty-related deficits such as cognition and nutrition. Each indicator received a weight based on its regression coefficient (in other words, its contribution to the likelihood of being frail). These weights were summed for each beneficiary creating the frailty score (Kim et al. 2019, Kim et al. 2020). The SAS code for the creation of the CFI developed by Kim et al. (2020) was provided from Harvard.edu. For each beneficiary, we measured the CFI as of the index date, one year prior to the index date, and two years prior to the index date. In each year, we used three CFI frailty categories: pre-frail (CFI < 0.25), mildly frail (CFI 0.25 to < 0.35), or moderately to severely frail (CFI ≥ 0.35) (Kim and Gautam 2020; Kim and Gautam 2024). The CFI has been validated in multiple cohorts and shown to predict outcomes such as death, hospital use, functional decline, and nursing home entry.
- We matched on indicators for the 10 most common modified diagnosis-related group (MDRG) categories from the beneficiary's most recent hospital admission, such as septicemia or severe sepsis (MDRG 1808).
- We matched on the number of months since the beneficiary was first diagnosed with each of the following: Alzheimer's disease, chronic obstructive pulmonary disease (COPD), heart failure, hip or pelvic fracture, or osteoporosis.
- We matched on hierarchical condition category (HCC) score not only in the year before index (as we did in the performance-year analysis) but also in months 13 to 24 and 25 to 36 before the index date. In addition, we matched on a select list of HCC indicators in months 13 to 24 before the index date.

- We matched on the change in total, inpatient, SNF, and home health spending from months 13 to 24 prior to the index date to months 1 to 12 prior to the index date. The reason we matched on the change in spending was to ensure that IAH and comparison beneficiaries were on similar health trajectories before their index dates. Although we did not use levels of total, inpatient, SNF, and home health spending within a year as matching variables, we examined balance on the levels of spending as part of selecting the final matched comparison group. We did this to increase the likelihood of selecting a matched comparison group that had parallel trends in spending.

As in the performance-year analysis, we matched each IAH beneficiary to up to five comparison beneficiaries. On average, each IAH beneficiary matched to 4.8 comparison beneficiaries.

3.3. Health status, functional status, and demographic characteristics for IAH and matched comparison beneficiaries

As in the performance-year analysis, we assessed the similarity of the IAH and comparison groups using the standardized difference in means, described in detail in Appendix C. We applied the same standards as in the performance-year analysis, achieving standardized differences less than 0.15 on all variables and less than 0.10 on the vast majority of variables. Exhibit D.3 describes the characteristics of the potential comparison group—the beneficiaries considered in the matching process—as well as the actual matched comparison group and the IAH group.

Exhibit D.3. Characteristics of IAH and matched comparison beneficiaries

Variable	Potential comparison group	Matched comparison group	IAH group	Standardized difference
Variables used in exact matching				
Number of months since last hospital admission, percentage				
One	41.3	48.8	48.8	0.000
Two or three	17.9	22.1	22.1	0.000
Four or more	40.7	29.0	29.0	0.000
Short-term death, percentage				
Beneficiary died within 3 months of index date	6.0	8.6	8.6	0.000
Demographic characteristics				
Age, percentage				
Younger than 65	9.6	7.4	7.6	0.007
65–79	38.5	33.3	29.4	-0.083
80 or older	51.9	59.3	63.0	0.075
Original reason for Medicare entitlement, percentage				
Eligibility due to age	76.0	77.8	78.5	0.018
Eligibility due to disability	22.0	21.0	20.5	-0.012
Eligibility due to ESRD or ESRD plus disability	2.0	1.2	1.0	-0.024

Exhibit D.3 (*continued*)

Variable	Potential comparison group	Matched comparison group	IAH group	Standardized difference
Dual eligibility for Medicare and Medicaid, percentage				
Non-dual	75.0	74.7	70.4	-0.097
Partial dual	5.1	4.2	4.3	0.006
Full dual	19.9	21.2	25.4	0.100
Number of ADLs requiring human assistance, percentage				
2 ADLs	14.6	9.4	7.5	-0.066
3-4 ADLs	34.6	32.5	29.7	-0.059
5-6 ADLs	50.8	58.2	62.7	0.093
Missing data on feeding ADL	9.0	10.2	12.3	0.066
Health status				
Observation stay without inpatient admission	0.06	0.037	0.049	0.059
HCC				
HCC risk score, months 1–12 before index	3.596	3.996	4.055	0.029
HCC risk score, months 13–24 before index	2.32	2.385	2.349	-0.022
HCC risk score, months 25–36 before index	2.055	2.071	2.046	-0.017
Percentage with HCCs in months 1–12 before index				
Immune deficiency (HCCs 2, 6, 46, 47, 186)	23.8	27.0	27.6	0.015
HCC8: Metastatic cancer	5.4	3.1	2.8	-0.016
HCC9-10: Lung and lymphoma cancers	6.7	5.4	4.7	-0.035
HCC11-12: Other cancers	9.0	8.6	7.5	-0.041
HCC18: Diabetes with complications	32.6	33.9	30.8	-0.066
HCC21: Protein-calorie malnutrition	12.9	16.7	20.6	0.101
HCC22: Morbid obesity	11.6	12.9	12.2	-0.019
HCC27: End-stage liver disease	1.6	1.6	1.4	-0.010
HCC28-29: Cirrhosis of liver and chronic hepatitis	2.4	2.0	1.9	-0.005
HCC46: Severe hematological disorders	2.2	1.8	1.6	-0.013
HCC48: Other hematological disorders	17.8	18.6	17.3	-0.034
HCC51: Dementia with complications	5.0	12.8	13.2	0.013
HCC52: Dementia without complications	15.5	21.3	27.0	0.133
HCC54-55: Drug/alcohol psychosis and drug/alcohol dependence	6.1	7.0	7.0	0.000
HCC57-58: Schizophrenia, major depressive, bipolar, and paranoid disorders	16.4	21.7	22.9	0.028
HCC70-71: Quadriplegia, paraplegia	2.1	2.9	4.4	0.080
HCC72: Spinal cord disorders/injuries	2.5	3.0	2.6	-0.025
HCC85: Congestive heart failure	49.1	53.4	50.2	-0.064

Exhibit D.3 (*continued*)

Variable	Potential comparison group	Matched comparison group	IAH group	Standardized difference
HCC96: Specified heart arrhythmias	43.2	47.3	44.3	-0.060
HCC103-104: Hemiplegia/hemiparesis, monoplegia, other paralytic syndromes	8.6	11.7	13.9	0.066
HCC106: Atherosclerosis of the extremities with ulceration or gangrene	3.5	3.8	3.8	-0.002
HCC107-108: Vascular disease	42.5	45.0	41.6	-0.069
HCC111: Chronic obstructive pulmonary disease	35.4	36.3	33.7	-0.055
HCC134: Dialysis status	5.5	4.1	3.3	-0.043
HCC136-138: Chronic kidney disease, stage 3-5	9.6	9.1	8.2	-0.032
HCC139-140: Chronic kidney disease, stage 1-2, unspecified renal failure	5.4	5.2	5.3	0.001
HCC157-159: Pressure ulcer	6.6	10.5	15.2	0.139
Percentage with HCCs in months 13–24 before index				
HCC103-104: Hemiplegia/hemiparesis, monoplegia, other paralytic syndromes	3.7	4.7	5.5	0.033
HCC18: Diabetes with complications	26.8	27.8	24.5	-0.075
HCC27: End-stage liver disease	0.9	0.7	0.5	-0.018
HCC28-29: Cirrhosis of liver and chronic hepatitis	1.8	1.3	1.2	-0.012
HCC57-58: Schizophrenia, major depressive, bipolar, and paranoid disorders	11.3	12.5	12.4	-0.004
HCC70-71: Quadriplegia, paraplegia	1.1	1.3	1.8	0.040
HCC134: Dialysis status	4.1	3.0	2.4	-0.036
HCC157-159: Pressure ulcer	2.4	3.1	4.2	0.058
Kim et al. frailty score, months 1–12 before index, percentage				
Pre-frail (CFI < 0.25)	30.8	11.8	11.6	-0.005
Mildly frail (CFI 0.25 to < 0.35)	46.1	43.8	40.9	-0.059
Moderately to severely frail (CFI ≥ 0.35)	23.1	44.4	47.5	0.062
Kim et al. frailty score, months 13–24 before index, percentage				
Pre-frail (CFI < 0.25)	65.0	56.8	57.0	0.005
Mildly frail (CFI 0.25 to < 0.35)	26.2	31.1	29.2	-0.040
Moderately to severely frail (CFI ≥ 0.35)	8.7	12.1	13.7	0.047
Kim et al. frailty score, months 25–36 before index, percentage				
Pre-frail (CFI < 0.25)	70.5	65.0	65.9	0.018
Mildly frail (CFI 0.25 to < 0.35)	22.8	26.2	24.2	-0.045
Moderately to severely frail (CFI ≥ 0.35)	6.7	8.8	9.9	0.038

Exhibit D.3 (*continued*)

Variable	Potential comparison group	Matched comparison group	IAH group	Standardized difference
MS-DRG, percentage				
MS-DRG with major complicating condition or complicating condition	56.9	60.8	60.7	-0.003
MDRG 0114: Intracranial hemorrhage or cerebral infarction	2.9	3.6	4.4	0.042
MDRG 0404: Respiratory infections and inflammations	1.2	1.4	1.6	0.013
MDRG 0410: Simple pneumonia and pleurisy	3.0	3.0	2.7	-0.016
MDRG 0524: Heart failure and shock	5.6	6.5	5.5	-0.042
MDRG 0807: Major hip and knee joint replacement or reattachment of lower extremity	5.1	2.4	1.9	-0.034
MDRG 0811: Hip and femur procedures except major joint	2.4	3.2	3.3	0.006
MDRG 1008: Miscellaneous disorders of nutrition, metabolism, fluids, and electrolytes	1.8	2.1	2.2	0.006
MDRG 1110: Renal failure	2.9	3.5	3.2	-0.017
MDRG 1113: Kidney and urinary tract infections	3.0	4.8	5.6	0.040
MDRG 1808: Septicemia or severe sepsis	5.3	6.5	7.2	0.028
Number of chronic conditions measured by Chronic Conditions Warehouse, percentage				
Fewer than six	26.1	22.0	27.5	0.127
Six to nine	50.1	48.6	47.6	-0.018
10 or more	23.8	29.4	24.8	-0.103
Other measures of health status, percentage				
Anemia	18.5	20.5	20.2	-0.008
Depression	38.6	46.7	47.5	0.016
Electrolyte disorder	33.7	37.9	39.5	0.032
Chronically critically ill/medically complex	33.2	37.1	38.7	0.033
Time since diagnosis				
Number of months since first diagnosed with Alzheimer's disease, percentage				
Never	63.1	46.8	39.4	-0.150
Up to 12 months	10.8	17.7	20.2	0.063
12 or more months	26.1	35.5	40.4	0.102
Number of months since first diagnosed with COPD, percentage				
Never	49.0	48.4	51.5	0.062
Up to 12 months	5.5	5.5	5.3	-0.009
12 or more months	45.5	46.1	43.3	-0.058

Exhibit D.3 (*continued*)

Variable	Potential comparison group	Matched comparison group	IAH group	Standardized difference
Number of months since first diagnosed with heart failure, percentage				
Never	37.7	33.9	35.7	0.038
Up to 12 months	8.6	9.3	9.8	0.016
12 or more months	53.8	56.8	54.5	-0.046
Number of months since hip or pelvic fracture, percentage				
Never	88.6	84.7	83.8	-0.025
Up to 12 months	4.3	6.3	6.4	0.003
12 or more months	7.1	9.0	9.9	0.029
Number of months since first diagnosed with osteoporosis, percentage				
Never	63.6	60.2	61.2	0.021
Up to 12 months	3.1	3.3	3.2	-0.009
12 or more months	33.3	36.5	35.6	-0.018
Expenditures				
Medicare inpatient expenditures ($PBPM)[a]				
Months 1–12 before index	$2,067	$2,401	$2,471	0.024
Months 13–24 before index	$730	$755	$768	0.007
Medicare home health expenditures ($PBPM)[a]				
Months 1–12 before index	$377	$457	$472	0.036
Months 13–24 before index	$140	$168	$196	0.080
Medicare SNF expenditures ($PBPM)[a]				
Months 1–12 before index	$552	$1,042	$1,163	0.098
Months 13–24 before index	$142	$177	$255	0.119
Total Medicare expenditures ($PBPM)[a]				
Months 1–12 before index	$4,225	$5,099	$5,182	0.020
Months 13–24 before index	$1,834	$1,861	$1,859	-0.001
Change in inpatient expenditures ($PBPM)				
Months 13–24 prior to index to months 1–12 prior to index	$1,338	$1,646	$1,703	0.018
Change in home health expenditures ($PBPM)				
Months 13–24 prior to index to months 1–12 prior to index	$237	$289	$275	-0.034
Change in SNF expenditures ($PBPM)				
Months 13–24 prior to index to months 1–12 prior to index	$409	$865	$908	0.032
Change in total Medicare expenditures ($PBPM)				
Months 13–24 prior to index to months 1–12 prior to index	$2,391	$3,239	$3,323	0.020

Exhibit D.3 (*continued*)

Source: Medicare claims and enrollment data for 2010–2021 obtained from the Chronic Conditions Warehouse for IAH and matched comparison group beneficiaries in all IAH practices included in the evaluation.

Notes: The final sample sizes were 9,783 IAH beneficiaries and 47,073 matched comparison beneficiaries. The number of weighted matched comparison beneficiaries equaled the number of IAH beneficiaries.

[a] These variables were not included in matching but were included in balance checks to assess the similarity of health trajectories between IAH and comparison beneficiaries.

ADLs = activities of daily living; CFI = claims-based frailty index; COPD = chronic obstructive pulmonary disease; ESRD = end-stage renal disease; HCC = hierarchical condition category; MDRG = modified diagnosis-related group; MS-DRG = Medicare severity diagnosis-related group; PBPM = per beneficiary per month; SNF = skilled nursing facility.

3.4. Number of beneficiaries and eligible months

As described above, beneficiaries joined the three-year analysis sample on their index date. The two years leading up to the index date are the beneficiary's pre-period. We followed beneficiaries for three years after the index date unless they died or left Medicare FFS. Thus, the number of IAH and comparison beneficiaries begins to decrease in the second year after the index date and decreases further in the third year, as beneficiaries die or leave Medicare FFS.

Across the follow-up period, the average number of eligible months was slightly lower for IAH beneficiaries than for comparisons, at least in part because more IAH beneficiaries than comparisons died in each year (for further discussion of death in the three-year analysis, see Chapter 11 of the main report). As described in more detail in Appendix C.5, we accounted for differences across beneficiaries in the number of eligible months in a year as part of our weighting strategy.

Exhibit D.4 shows the number of beneficiaries, total number of eligible months, and average number of eligible months in each year, separately for the IAH and comparison groups.

Exhibit D.4. Sample for the three-year analysis, by period

	Two years before index	Year before index	First year	Second year	Third year
Number of IAH beneficiaries	9,783	9,783	9,783	6,552	4,294
Total number of eligible months for IAH beneficiaries	117,396	117,396	98,015	64,015	42,073
Average number of eligible months per IAH beneficiary	12.0	12.0	10.0	9.8	9.8
Number of comparison beneficiaries	47,073	47,073	47,073	33,599	23,847
Total number of eligible months for comparison beneficiaries	564,876	564,876	482,662	341,896	242,406
Average number of eligible months per comparison beneficiary	12.0	12.0	10.3	10.2	10.2

Source: Medicare claims and enrollment data for 2010–2022 obtained from the Chronic Conditions Warehouse for IAH and matched comparison group beneficiaries in all IAH practices included in the evaluation.

4. Methods for estimating impacts

As in the performance-year analysis, in the three-year analysis we used a difference-in-differences model to estimate the impact of IAH in the first, second, and third years after beneficiaries began receiving HBPC from an IAH practice.

4.1. Interpretation of estimated effects

The three-year analysis estimates the effect of newly receiving HBPC from an IAH practice during the IAH demonstration (Exhibit D.1). This interpretation differs from the interpretation of the estimated effects from the performance-year analysis, which represent the effects of the demonstration (that is, the opportunity to earn incentive payments) for new and existing users of HBPC (Exhibit C.1).

Both qualitative and quantitative evidence strongly indicate that HBPC drove the observed effects in the three-year analysis. Qualitatively, IAH practices responded to demonstration incentives by making only minor changes in care delivery, as we discuss in Chapter 1 of the report. Their changes were geared toward reducing hospital and ED use and becoming aware of their beneficiaries' hospital admissions and ED visits. Quantitatively, the three-year analysis tended to find much larger effects than the performance-year analysis. For example, where the performance-year analysis found a 4.9 percent reduction in total Medicare spending on average over 10 demonstration years, the three-year analysis found a 9.4 percent reduction in total Medicare spending in the first year after beneficiaries started HBPC. This juxtaposition implies that the component that differs between the two analyses—starting HBPC at an IAH practice—drives the difference in estimated effects. Additional sensitivity analyses supported this conclusion (Exhibit D.5).

4.2. Difference-in-differences model

4.2.1. Model specifications

For each outcome, we estimated the impact of newly receiving HBPC from an IAH practice during the demonstration with a difference-in-differences regression. This approach compares the difference in outcomes between a follow-up period and a reference period, among beneficiaries who received HBPC, to the difference in outcomes over the same period among beneficiaries who did not receive HBPC but were otherwise similar to those who did. In the three-year analysis of IAH beneficiaries new to HBPC, the reference period was the year before the index date and the follow-up periods were each of the three years after the index date.

The regression equation for this analysis is:

$$(1)\quad \begin{aligned} Y_{it} &= \alpha + X_{it}\beta + \tau Treatment_i + \gamma_1 P1 + \gamma_2 P3 + \gamma_3 P4 + \gamma_4 P5 + \theta_1 P1 \cdot Treatment_i \\ &+ \theta_2 P3 \cdot Treatment_i + \theta_3 P4 \cdot Treatment_i + \theta_4 P5 \cdot Treatment_i + \varepsilon_{it} \end{aligned}$$

In this equation, Y_{it} is the outcome variable for beneficiary *i* measured in period *t*, α is the overall intercept, and β adjusts for beneficiary characteristics X_{it} measured contemporaneously with outcomes. The remaining terms estimate the difference-in-differences: τ represents the difference in

outcomes between IAH and comparison beneficiaries in the reference year, with treatment status indicated by $Treatment_i$; the γ_t terms represent the difference in outcomes in the comparison group between the reference year and each of four non-reference periods, from the year before the reference year (P1) through three years after the reference year (P5); and the θ_t terms represent the difference-in-differences in each non-reference period. ε_{it} is a residual error term.

We estimate a separate θ_t term in each follow-up period to allow for nonlinear trends in the estimated effects of the intervention. At the same time, we estimate θ_1, the difference-in-differences two years before the beneficiary began receiving HBPC, as a test of our modeling assumptions; we discuss this test later in the chapter.

We followed the conventions of the performance-year analysis when determining the functional form of the regression for each claims-based outcome and, in consequence, when transforming the difference-in-difference coefficients θ_t into estimated effects. As in the performance-year analysis, we used more complex specifications to estimate the intervention's effects on death, the probability of an unplanned hospital readmission, and the probability of entry into a nursing home. We describe these specifications later in the chapter.

We translated estimated effects into percentage terms by dividing each estimated effect by the adjusted IAH group mean for the same outcome in the year before beneficiaries started HBPC.

As in the performance-year analysis, we adjusted standard errors for clustering using the hybrid approach and used beneficiary weighting to account for matching and observability (see Appendix C.5). We describe these approaches in more detail in later sections of the chapter.

4.2.2. Model specification for death

The three-year analysis has a retrospective cohort design: beneficiaries join the sample on the date of their first HBPC visit and we measure their outcomes for three years after this date, as well as two years before this date. By construction, then, all members of our sample, in both the IAH and comparison groups, are alive in the year before they join the sample. As a result, we cannot use our primary difference-in-differences approach to estimate the intervention's effect on death.

Instead, we used the survival modeling techniques used in the performance-year analysis, with one key difference: rather than fitting a difference-in-difference regression in this framework, we estimated an IAH-comparison difference in the follow-up period. The regression equation was:

(2) $\log\left(T_i\right) = X_i\beta + \tau Treatment_i + z_i$

where T_i denotes the number of days that beneficiary *i* survived after joining the sample; X_i contains the same set of covariates included in Equation (1), measured in the reference year; $Treatment_i$ is an

indicator for whether the beneficiary was in the IAH group; and z_i is an error term with a generalized gamma distribution.[45]

The model specification in Equation (2) accounts for beneficiaries whose survival time was not observed—for example, because they were still alive at the end of follow-up. We estimated the model using the maximum likelihood method. We used matching weights to account for the number of matched comparisons per IAH beneficiary, ensuring balance on covariates and equality in sample size in the two groups. We adjusted standard errors using the hybrid clustering approach, described in more detail in Appendix C.5.

From the fitted model we generated survival probabilities for each group, IAH and comparison, in each follow-up year, and converted these survival probabilities to death rates by subtracting the survival probabilities from one. The estimated effects are the differences in estimated death rates, IAH group versus comparison group, in the first, second, and third years since starting HBPC.

This cross-sectional approach is not as rigorous as the difference-in-differences strategy used to estimate effects for other claims-based outcomes. Specifically, with a cross-sectional approach we cannot account for differences between the IAH and comparison groups that predate the intervention; if IAH beneficiaries tend to prefer less aggressive end-of-life care, for example, they may have higher death rates than comparison beneficiaries, even without receiving HBPC during the demonstration. Thus, because we only observe death in the follow-up period, we cannot ascertain how much of the difference in death rates reflects the characteristics and preferences of the IAH and comparison beneficiaries, versus the influence of starting HBPC during the demonstration.

4.2.3. Model specification for other binary outcomes

In addition to death, we estimated the effect of newly receiving HBPC from an IAH practice on two other binary outcomes: the probability of an unplanned readmission and the probability of entering a nursing home within a given follow-up year. We used a similar approach to the performance-year analysis when estimating effects for these outcomes, but the approach differed in several important ways.

As in the performance-year analysis, our overall strategy was to estimate a difference-in-differences logistic regression with the same form as Equation (1) and to account for differential eligibility between IAH and comparison beneficiaries by controlling for the predicted probability of death in each period. (Unlike in the performance-year analysis, IAH beneficiaries tended to be eligible for less of the follow-up period than comparison beneficiaries; for example, in the first follow-up year 28 percent of IAH beneficiaries new to HBPC died, compared to 25 percent of comparison beneficiaries.)

However, because all members of the three-year analysis sample are alive until the first follow-up year, we cannot use Equation (2) to predict a probability of death in the pre-period. Instead, for the three-year analysis we computed the predicted probability of death using a logistic regression model fit to data from the comparison group in the first follow-up year. The predictors for this model included all the predictors X_i from the main regression described in Appendix D.4.2.1. We used the predictions

[45] Although we could have considered alternative distributions for the survival function, we chose to use a generalized gamma distribution for consistency with the performance-year study.

from this model for all beneficiaries, IAH and comparison, in all periods, including both the pre-period and the three follow-up periods.

This strategy allows us to adjust for differences between the IAH and comparison groups in risk factors associated with death. However, unlike the approach used in the performance-year analysis, it does not allow us to adjust for effects of HBPC from IAH practices on death. As a result, estimates from the three-year analysis of effects on binary outcomes (such as the probability of entering a nursing home) also include any effects of IAH on the probability of death.

Equation (3) gives our model specification for non-death binary outcomes in the three-year analysis.

$$(3)\quad \begin{aligned} P\left(Y_{it}=1\right) &= \alpha + X_{it}\beta + \tau Treatment_{it} + \beta_p p_{it} + \gamma_1 P1 + \gamma_2 P3 + \gamma_3 P4 + \gamma_4 P5 + \theta_1 P1 \cdot Treatment_i \\ &+ \theta_2 P3 \cdot Treatment_i + \theta_3 P4 \cdot Treatment_i + \theta_4 P5 \cdot Treatment_i \end{aligned}$$

In this equation, p_{it} represents the predicted probability that beneficiary *i* dies at time *t*; all other terms also appear in Equation (1). We estimated the model using logistic regression, adjusting for matching weights and applying the hybrid clustering approach, as in other regressions. As in the performance-year analysis, we used a bootstrap procedure to account for the fact that p_{it} is itself an estimate derived from a separate regression model. After bootstrapping, we transformed the coefficients $\theta_1, \theta_2, \ldots, \theta_4$ into difference-in-differences effect estimates following steps similar to those used to estimate effects for other outcomes.

4.2.4. Weighting and adjustment to standard errors for clustering

The three-year analysis followed the same weighting and hybrid clustering approach as the performance-year analysis, described in full in Appendix C.5.

4.2.5. Control variables

As in the performance-year analysis, although the comparison group selected for the three-year analysis was well balanced on a range of variables, we controlled for a comprehensive set of variables in our regressions, to adjust for any remaining imbalance and to increase the precision of the impact estimates. These variables fall into four main categories: (1) variables describing eligibility for the demonstration; (2) demographic and enrollment characteristics; (3) ADL indicators; and (4) measures of health status, including claims-based frailty index, HCC risk score, HCC indicators, and chronic condition indicators. We included all HCC indicators and categories of HCCs used for matching (Exhibit D.3), but some of the other control variables were included in regressions at a more detailed level than the variables we used in matching. For example, we used three age categories in propensity score matching, whereas we used five age categories in the outcome regressions.

Many of these variables are the same as the variables used as controls in the performance-year analysis, with two key differences. First, because the three-year analysis sample comprises beneficiaries who started HBPC between 2013 and 2021, we controlled for the calendar year in which the beneficiary joined the sample. Second, we included a control for the combination of the state in which the IAH beneficiary resided and the IAH practice from which they received care. We did not include an IAH

practice control in the performance-year analysis because, when combined with the hybrid clustering approach (see Appendix C.5), it could have produced counterintuitive results. The same concern did not apply to the three-year analysis because it estimated effects across all demonstration years, rather than separately by demonstration year.

4.2.6. Testing the parallel trends assumption

As described in more detail in Chapter 5 of this appendix, the difference-in-differences regression approach used to estimate the effect of newly receiving HBPC from an IAH practice is not valid if the IAH and matched comparison groups exhibited substantially different outcome trends in the pre-period. In the three-year analysis we adopted the same strategy as in the performance-year analysis, using the estimated effect two years before beneficiaries joined the sample to gauge the difference in pre-period trends between the IAH and matched comparison groups.

The estimated effects two years before beneficiaries joined the sample were neither statistically significant at the 90 percent level nor substantial (larger than 5 percent of the reference year mean in the IAH group) for our primary outcomes of total Medicare spending, inpatient spending, SNF spending, and hospital admissions. However, we did find statistically significant differences in pre-period trends for some other outcomes. These outcomes fall into two categories:

1. The difference in trends does not pose a threat to our estimated effects because, if the pre-period trends had continued into the follow-up period, the estimated effects would have had the opposite sign. For example, DME spending was increasing faster in the comparison group than in the IAH group in the pre-period. If these trends had continued into the follow-up period, we would expect to see that DME spending decreased among IAH beneficiaries, relative to matched comparisons. Instead, the opposite was true: we found that DME spending increased more among IAH beneficiaries than comparisons.
2. While statistically significant, the difference in trends is small compared to the estimated effects. For example, hospice spending increased by $28 per beneficiary per month (PBPM) more among IAH beneficiaries than comparisons in the year before starting HBPC, but we estimate increases in hospice spending of between $150 and $200 PBPM. Even if we assume that pre-existing trends continued unchanged into the follow-up period, they would explain less than half of the estimated effect in the third year after beneficiaries started HBPC ($84 of $195 PBPM, or 43 percent).

4.2.7. Sensitivity analyses

To test the robustness of our conclusions to assumptions of our data construction process and model, we conducted several sensitivity analyses.

The first sensitivity analysis, discussed in Chapter 9 of the report, compared beneficiaries who started HBPC in 2013 or 2014 to those who started in 2015 or later. This sensitivity analysis took advantage of IAH practices' gradual acclimation to the demonstration. We know from interviews that it took time for IAH practices to decide how to respond to the opportunity to earn incentive payments. Therefore, it is unlikely that the minor changes made by IAH practices because of participating in the demonstration affected outcomes for beneficiaries who started HBPC early in the demonstration in the same way as they may have affected outcomes for beneficiaries who started HBPC later on, when the demonstration was well underway. If the demonstration drove the estimated effects in the three-year analysis, we would expect to see larger estimated effects in later years of the demonstration. Yet estimated effects

from the three-year analysis were very similar for beneficiaries who started HBPC at an IAH practice in the first two years of the demonstration, 2013–2014, and those who started HBPC in 2015–2021 (Exhibit D.5).

Exhibit D.5. Estimated effects were very similar for beneficiaries who started HBPC at an IAH practice in 2013–2014 and 2015 or later

Period	Started HBPC in 2013–2014 (early in the demonstration)		Started HBPC in 2015–2021 (later in the demonstration)		Difference in estimated effect early versus late
	Difference-in-differences estimated effect (SE)	Percentage effect	Difference-in-differences estimated effect (SE)	Percentage effect	
Third year	-100 (207)	-2.0	-55 (221)	-1.0	45 (166)
Second year	-15 (136)	-0.3	16 (186)	0.3	31 (191)
First year	-548*** (95)	-11	-481*** (156)	-8.9	67 (138)
One year pre-index	n.a.	n.a.	n.a.	n.a.	n.a.
Two years pre-index	-155 (118)	-3.1	-77 (199)	-1.4	78 (170)

Source: Medicare claims and enrollment data for 2010–2022 obtained from the Chronic Conditions Warehouse for IAH and matched comparison group beneficiaries in all IAH practices included in the evaluation.

Notes: Standard errors are in parentheses. We computed coefficients and standard errors by using the weighted sample size, which considers the matching and eligibility weights. Because of rounding, a difference-in-differences estimate displayed as zero might be shown alongside a percentage effect that exceeds zero.

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

HBPC = home-based primary care; n.a. = not applicable; SE = standard error.

In other words, the very large reduction in spending in the first year was present immediately and consistent over the entire study period. The immediacy and magnitude of the change suggests that effects of the demonstration were caused mostly by starting HBPC, not by the minor changes IAH practices made because of participating in the demonstration.

We conducted two sensitivity analyses that assess how we implemented our intent-to-treat design. First, we assessed the changes to our results if we required members of the intervention group to receive HBPC from an IAH practice in all three follow-up years, rather than in the first year only. Second, we gauged the effect of measuring IAH beneficiaries' outcomes only while their IAH practice was participating in the demonstration.

In the first of these analyses, we estimated effects on total Medicare spending in the subset of beneficiaries who continued to receive HBPC from their IAH practice in the second and third follow-up years because effects among the full sample in the second and third years could have been diluted by beneficiaries who were not actively receiving HBPC from an IAH practice. Nonetheless, estimated effects when limiting the intervention group to beneficiaries who continued to receive HBPC from an IAH practice were not statistically significantly different from the main results in the second and third follow-up years, as Exhibit D.6 shows. Larger estimated effects in the second and third years could reflect heightened selection, if beneficiaries who continued receiving HBPC were those who found it particularly effective.

Exhibit D.6. Estimated effects on total Medicare spending among beneficiaries who continued to receive HBPC one, two, and three years after starting HBPC

Period	IAH	Comparison	Alternative difference-in-differences estimated effect (SE)	Main difference-in-differences estimated effect (SE)	Percentage effect[a]
Third year	$2,843	$2,943	-$299 ($338)	-$68 ($204)	-5.6
Second year	$3,163	$3,044	-$80 ($215)	$7 ($151)	-1.5
First year	$4,007	$4,311	-$503*** ($128)	-$501*** ($127)	-9.5
One year pre-index	$5,300	$5,101	-	-	-
Two years pre-index	$2,042	$1,942	-$99 ($163)	-$99 ($164)	-1.9

Source: Medicare claims and enrollment data for 2010–2022 obtained from the Chronic Conditions Warehouse for IAH and matched comparison group beneficiaries in all IAH practices included in the evaluation.

Notes: Standard errors are in parentheses. We computed coefficients and standard errors by using the weighted sample size, which considers the matching and eligibility weights. Because of rounding, a difference-in-differences estimate displayed as zero might be shown alongside a percentage effect that exceeds zero.

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

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

SE = standard error.

In the second of these analyses, we determined how much estimated effects changed if we excluded the time after a beneficiary's practice exited the demonstration (Exhibit D.7). For this sensitivity analysis, we measured a beneficiary's outcomes during a given follow-up year if the practice remained in the demonstration for at least six months of that follow-up year (that is, we applied a majority rule). For example, the second follow-up year for an IAH beneficiary at the Portland practice with an index date of June 1, 2019, would begin on June 1, 2020, and end on May 31, 2021. Because seven months of the beneficiary's second follow-up year took place while Portland was participating in the intervention (June through December 2020), we would have considered this beneficiary's second follow-up year as part of the intervention period for the sensitivity analysis, but not their third follow-up year.

Limiting to beneficiaries whose practice still participated in the demonstration, regardless of whether they continued to receive HBPC, produced similar results as the main analysis.

Exhibit D.7. Estimated effects on total Medicare spending among beneficiaries whose IAH practice remained in the demonstration, one, two, and three years after starting HBPC

Period	IAH	Comparison	Difference-in-differences estimated effect (SE)	Main difference-in-differences estimated effect (SE)	Percentage effect[a]
Third year	$3,040	$2,964	-$125 ($208)	-$68 ($204)	-2.3
Second year	$3,297	$3,082	$13 ($149)	$7 ($151)	0.2
First year	$4,048	$4,349	-$502*** ($127)	-$501*** ($127)	-9.4
One year pre-index	$5,345	$5,144	-	-	-
Two years pre-index	$2,089	$1,987	-$99 ($163)	-$99 ($164)	-1.9

Source: Medicare claims and enrollment data for 2010–2022 obtained from the Chronic Conditions Warehouse for IAH and matched comparison group beneficiaries in all IAH practices included in the evaluation.

Notes: Standard errors are in parentheses. We computed coefficients and standard errors by using the weighted sample size, which considers the matching and eligibility weights. Because of rounding, a difference-in-differences estimate displayed as zero might be shown alongside a percentage effect that exceeds zero.

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

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

SE = standard error.

Next we conducted an analysis to gauge the influence each IAH practice has on the estimated effects. As in the performance-year analysis (Appendix C.5), we assessed the practices' influence by removing each practice from the sample one at a time and refitting our regression models to the modified sample. If the estimated effects excluding any particular practice differed dramatically from the estimated effects based on the full sample, that practice was very influential. Influential practices may lower our level of confidence in the full sample results because we cannot be sure that the results we observed were caused by starting HBPC during the demonstration generally, rather than something unique about specific practices. In general, we found that removing individual practices did not change our conclusions, which strengthens our confidence in the results.

Finally, we tested for the influence of outlier or high spending observations at the beneficiary level (Exhibit D.23). To do so, we computed the 99th percentile of the distribution of spending among IAH beneficiaries, separately in each calendar year. We winsorized total Medicare spending by setting any values higher than the 99th percentile to the 99th percentile value among IAH beneficiaries in that calendar year. We then refit our regressions using the winsorized data; differences between the original and winsorized estimates reflect the influence of outlying values. We also found very little difference before and after winsorizing, suggesting that effects were not driven by very high spending beneficiaries.

4.2.8. Subgroup analyses

To better understand which groups of beneficiaries benefit most from the specialized HBPC that IAH practices provide, we estimated effects for subgroups of beneficiaries who may have been at higher risk of poor outcomes and high spending as a result of functional impairment or social disadvantage. 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 (2-4), (2) beneficiaries with dual Medicare and Medicaid coverage versus beneficiaries with Medicare alone, and (3) all combinations of these two subgroups (for example, dually eligible beneficiaries who require assistance with 5 or 6 ADLs). Exhibit D.8 gives sample sizes for all subgroups.

Exhibit D.8. Sample sizes by subgroup for IAH and comparison groups

Subgroup	IAH beneficiaries (percentage)	Comparison beneficiaries (percentage)
Dual eligibility for Medicare and Medicaid		
Non-dually eligible	6,883 (70.4%)	35,231 (74.8%)
Dually eligible	2,900 (29.6%)	11,842 (25.2%)
ADLs		
Required assistance from another person with 2–4 ADLs	3,645 (37.3%)	19,845 (42.2%)
Required assistance from another person with 5 or 6 ADLs	6,138 (62.7%)	27,228 (57.8%)
Combination of dual eligibility for Medicare and Medicaid and functional impairment		
Non-dually eligible, required assistance with 2–4 ADLs	2,485 (25.4%)	14,591 (31.0%)
Non-dually eligible, required assistance with 5 or 6 ADLs	4,398 (45.0%)	20,640 (43.8%)
Dually eligible, required assistance with 2–4 ADLs	1,160 (11.9%)	5,254 (11.2%)
Dually eligible, required assistance with 5 or 6 ADLs	1,740 (17.8%)	6,588 (14.0%)

Source: Medicare claims and enrollment data for 2010–2022 obtained from the Chronic Conditions Warehouse for IAH and matched comparison group beneficiaries in all IAH practices included in the evaluation.

ADLs = activities of daily living.

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 since starting HBPC at an IAH practice). 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 for one group (for example, beneficiaries with dual Medicare and Medicaid coverage) were different from the other group (for example, beneficiaries with Medicare alone) in each year since beneficiaries began receiving HBPC. All spending reported for beneficiaries dually eligible for Medicare and Medicaid is Medicare spending only.

5. Supplementary tables

Exhibit D.9. Estimated effect on primary care visits, one, two, and three years after starting HBPC

Period	IAH	Comparison	Difference (IAH – comparison)	Difference-in-differences estimated effect (SE)	Percentage effect[a]
Third year	8.0	5.1	2.9	5.0*** (0.4)	104.7
Second year	8.8	5.2	3.6	5.7*** (0.5)	119.9
First year	11.3	6.1	5.2	7.3*** (0.7)	154.1
One year pre-index	4.7	6.8	-2.1	-	-
Two years pre-index	5.1	6.4	-1.3	0.8*** (0.1)	16.6

Source: Medicare claims and enrollment data for 2010–2022 obtained from the Chronic Conditions Warehouse for IAH and matched comparison group beneficiaries in all IAH practices included in the evaluation.

Notes: Standard errors are in parentheses. We computed coefficients and standard errors by using the weighted sample size, which considers the matching and eligibility weights. Because of rounding, a difference-in-differences estimate displayed as zero might be shown alongside a percentage effect that exceeds zero.

Primary care visits include visits at home and visits by telehealth and telephone, as well as those in offices and clinics.

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

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

SE = standard error.

Exhibit D.10. Estimated effect on specialty care visits, one, two, and three years after starting HBPC

Period	IAH	Comparison	Difference (IAH – comparison)	Difference-in-differences estimated effect (SE)	Percentage effect[a]
Third year	3.1	5.0	-2.0	0.4* (0.2)	7.9
Second year	3.3	5.5	-2.1	0.2 (0.2)	4.4
First year	3.9	6.8	-2.9	-0.5*** (0.1)	-9.4
One year pre-index	5.6	8.0	-2.4	-	-
Two years pre-index	5.9	7.5	-1.7	0.7*** (0.1)	12.7

Source: Medicare claims and enrollment data for 2010–2022 obtained from the Chronic Conditions Warehouse for IAH and matched comparison group beneficiaries in all IAH practices included in the evaluation.

Notes: Standard errors are in parentheses. We computed coefficients and standard errors by using the weighted sample size, which considers the matching and eligibility weights. Because of rounding, a difference-in-differences estimate displayed as zero might be shown alongside a percentage effect that exceeds zero.

Specialty care visits include visits at home and visits by telehealth and telephone, as well as those in offices and clinics.

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

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

SE = standard error.

Exhibit D.11. Estimated effect on advanced imaging spending, one, two, and three years after starting HBPC

Period	IAH	Comparison	Difference (IAH – comparison)	Difference-in-differences estimated effect (SE)	Percentage effect[a]
Third year	$166	$246	-$80	-$9 ($11)	-2.0
Second year	$201	$279	-$78	-$7 ($12)	-1.6
First year	$262	$379	-$117	-$46*** ($15)	-10.2
One year pre-index	$450	$521	-$71	-	-
Two years pre-index	$291	$347	-$56	$15** ($6)	3.3

Source: Medicare claims and enrollment data for 2010–2022 obtained from the Chronic Conditions Warehouse for IAH and matched comparison group beneficiaries in all IAH practices included in the evaluation.

Notes Standard errors are in parentheses. We computed coefficients and standard errors by using the weighted sample size, which considers the matching and eligibility weights. Because of rounding, a difference-in-differences estimate displayed as zero might be shown alongside a percentage effect that exceeds zero.

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

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

SE = standard error.

Exhibit D.12. Estimated effect on testing spending, one, two, and three years after starting HBPC

Period	IAH	Comparison	Difference (IAH – comparison)	Difference-in-differences estimated effect (SE)	Percentage effect[a]
Third year	$595	$508	$87	$225*** ($63)	39.1
Second year	$621	$528	$93	$232*** ($55)	40.2
First year	$808	$591	$217	$356*** ($60)	61.7
One year pre-index	$577	$715	-$139	-	-
Two years pre-index	$525	$622	-$97	$42*** ($10)	7.2

Source: Medicare claims and enrollment data for 2010–2022 obtained from the Chronic Conditions Warehouse for IAH and matched comparison group beneficiaries in all IAH practices included in the evaluation.

Notes: Standard errors are in parentheses. We computed coefficients and standard errors by using the weighted sample size, which considers the matching and eligibility weights. Because of rounding, a difference-in-differences estimate displayed as zero might be shown alongside a percentage effect that exceeds zero.

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

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

SE = standard error.

Exhibit D.13. Estimated effect on other imaging spending, one, two, and three years after starting HBPC

Period	IAH	Comparison	Difference (IAH – comparison)	Difference-in-differences estimated effect (SE)	Percentage effect[a]
Third year	$360	$373	-$13	$100*** ($22)	19.8
Second year	$402	$394	$8	$121*** ($22)	23.8
First year	$516	$492	$23	$136*** ($30)	26.9
One year pre-index	$508	$621	-$113	-	-
Two years pre-index	$366	$450	-$84	$29* ($16)	5.6

Source: Medicare claims and enrollment data for 2010–2022 obtained from the Chronic Conditions Warehouse for IAH and matched comparison group beneficiaries in all IAH practices included in the evaluation.

Notes: Standard errors are in parentheses. We computed coefficients and standard errors by using the weighted sample size, which considers the matching and eligibility weights. Because of rounding, a difference-in-differences estimate displayed as zero might be shown alongside a percentage effect that exceeds zero.

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

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

SE = standard error.

Exhibit D.14. Estimated effect on receiving chronic care management services, one, two, and three years after starting HBPC

Period	IAH	Comparison	Difference (IAH – comparison)	Difference-in-differences estimated effect (SE)	Percentage effect[a]
Third year	28.26	10.07	18.19	18.58*** (4.11)	600.9
Second year	23.74	6.04	17.69	18.08*** (2.23)	584.9
First year	25.25	3.75	21.50	21.89*** (2.48)	708.1
One year pre-index	3.09	3.48	-0.39	-	-
Two years pre-index	2.23	2.83	-0.60	-0.22 (0.35)	-7.0

Source: Medicare claims and enrollment data for 2010–2022 obtained from the Chronic Conditions Warehouse for IAH and matched comparison group beneficiaries in all IAH practices included in the evaluation.

Notes: CMS began reimbursing health care providers for chronic care management services in 2014. Therefore, the analysis is limited to beneficiaries who began receiving HBPC from an IAH practice in 2016 or later, so that they could have received chronic care management services both before and after starting HBPC.

Standard errors are in parentheses. We computed coefficients and standard errors by using the weighted sample size, which considers the matching and eligibility weights. Because of rounding, a difference-in-differences estimate displayed as zero might be shown alongside a percentage effect that exceeds zero.

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

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

CMS = Centers for Medicare & Medicaid Services; SE = standard error.

Exhibit D.15. Estimated effect on outpatient ED visits and potentially avoidable outpatient ED visits, one, two, and three years after starting HBPC

Period	IAH	Comparison	Difference (IAH – comparison)	Difference-in-differences estimated effect (SE)	Percentage effect[a]
Outpatient ED visits per 1,000 beneficiaries per year					
Third year	900	1,021	-121	-211*** (53)	-12.1
Second year	1,027	1,100	-72	-163*** (56)	-9.3
First year	1,292	1,314	-22	-113 (87)	-6.5
One year pre-index	1,747	1,656	91	-	-
Two years pre-index	1,046	1,068	-22	-113*** (25)	-6.5
Potentially avoidable outpatient ED visits per 1,000 beneficiaries per year					
Third year	124	136	-13	-35*** (13)	-16.3
Second year	127	139	-12	-35*** (12)	-16.0
First year	156	164	-7	-30** (13)	-13.8
One year pre-index	218	195	23	-	-
Two years pre-index	130	128	2	-21** (9)	-9.6

Source: Medicare claims and enrollment data for 2010–2022 obtained from the Chronic Conditions Warehouse for IAH and matched comparison group beneficiaries in all IAH practices included in the evaluation.

Notes: Standard errors are in parentheses. We computed coefficients and standard errors by using the weighted sample size, which considers the matching and eligibility weights. Because of rounding, a difference-in-differences estimate displayed as zero might be shown alongside a percentage effect that exceeds zero.

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

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

ED = emergency department; SE = standard error.

Exhibit D.16. Estimated effect on hospital admissions and potentially avoidable hospital admissions, one, two, and three years after starting HBPC

Period	IAH	Comparison	Difference (IAH – comparison)	Difference-in-differences estimated effect (SE)	Percentage effect[a]
Hospital admissions per 1,000 beneficiaries per year					
Third year	1,101	1,212	-111	-80 (52)	-3.2
Second year	1,128	1,282	-154	-123** (58)	-5.0
First year	1,489	1,803	-314	-283*** (40)	-11.4
One year pre-index	2,475	2,506	-31	-	-
Two years pre-index	867	867	-1	31 (42)	1.2
Potentially avoidable hospital admissions per 1,000 beneficiaries per year					
Third year	252	276	-24	-17 (14)	-3.7
Second year	236	282	-46	-39*** (13)	-8.5
First year	314	372	-58	-50*** (13)	-11.1
One year pre-index	452	460	-7	-	-
Two years pre-index	197	183	14	21** (9)	4.7
Hospital admissions preceded by ED visit per 1,000 beneficiaries per year					
Third year	917	974	-57	-87** (43)	-4.5
Second year	920	999	-79	-109** (55)	-5.6
First year	1,199	1,380	-181	-211*** (31)	-10.9
One year pre-index	1,945	1,915	30	-	-
Two years pre-index	638	602	36	5 (41)	0.3

Source: Medicare claims and enrollment data for 2010–2022 obtained from the Chronic Conditions Warehouse for IAH and matched comparison group beneficiaries in all IAH practices included in the evaluation.

Notes: Standard errors are in parentheses. We computed coefficients and standard errors by using the weighted sample size, which considers the matching and eligibility weights. Because of rounding, a difference-in-differences estimate displayed as zero might be shown alongside a percentage effect that exceeds zero.

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

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

ED = emergency department; SE = standard error.

Exhibit D.17. Estimated effect on the probability of having a qualifying hospital discharge and unplanned readmission within 30 days of discharge, one, two, and three years after starting HBPC

Period	IAH	Comparison	Difference (IAH – comparison)	Difference-in-differences estimated effect (SE)	Percentage effect[a]
Third year	10.50	11.84	-1.34	-2.00* (1.12)	-7.9
Second year	11.25	12.41	-1.16	-1.83* (0.97)	-7.2
First year	14.77	19.15	-4.38	-5.05*** (0.83)	-19.8
One year pre-index	25.45	24.79	0.67	-	-
Two years pre-index	8.03	7.82	0.22	-0.45 (0.80)	-1.8

Source: Medicare claims and enrollment data for 2010–2022 obtained from the Chronic Conditions Warehouse for IAH and matched comparison group beneficiaries in all IAH practices included in the evaluation.

Notes: Standard errors are in parentheses. We computed coefficients and standard errors by using the weighted sample size, which considers the matching and eligibility weights. Because of rounding, a difference-in-differences estimate displayed as zero might be shown alongside a percentage effect that exceeds zero.

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

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

SE = standard error.

Exhibit D.18. Estimated effect on total Medicare spending, one, two, and three years after starting HBPC

Period	IAH	Comparison	Difference IAH – comparison (SE)	Difference-in-differences estimated effect (SE)	Percentage effect[a]
Third year	$3,083	$2,952	$132 ($120)	-$68 ($204)	-1.3
Second year	$3,283	$3,076	$208 ($68)	$7 ($151)	0.1
First year	$4,041	$4,342	-$301 ($49)	-$501*** ($127)	-9.4
One year pre-index	$5,341	$5,140	$200 ($104)	-	-
Two years pre-index	$2,088	$1,987	$101 ($66)	-$99 ($164)	-1.9

Source: Medicare claims and enrollment data for 2010–2022 obtained from the Chronic Conditions Warehouse for IAH and matched comparison group beneficiaries in all IAH practices included in the evaluation.

Notes: Standard errors are in parentheses. We computed coefficients and standard errors by using the weighted sample size, which considers the matching and eligibility weights. Because of rounding, a difference-in-differences estimate displayed as zero might be shown alongside a percentage effect that exceeds zero.

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

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

SE = standard error.

Exhibit D.19. Estimated effect on total Medicare spending by level of functional impairment, one, two, and three years after starting HBPC

Period	Needed help with 2–4 ADLs				Needed help with 5–6 ADLs				Difference in estimated effects, subgroup vs. reference (SE)
	IAH	Comparison	Difference-in-differences estimated effect (SE)	Percentage effect[a]	IAH	Comparison	Difference-in-differences estimated effect (SE)	Percentage effect[a]	
Third year	$2,944	$2,683	$436** ($207)	9.1	$3,080	$3,075	-$471** ($210)	-8.3	-$907*** ($156)
Second year	$2,877	$2,704	$348** ($155)	7.3	$3,548	$3,307	-$235 ($172)	-4.1	-$583*** ($153)
First year	$2,845	$2,939	$81 ($200)	1.7	$4,853	$5,409	-$1,032*** ($126)	-18.1	-$1,113*** ($167)
One year pre-index	$4,799	$4,974	-	-	$5,698	$5,223	-	-	-
Two years pre-index	$1,856	$1,795	$236 ($161)	4.9	$2,261	$2,085	-$300* ($168)	-5.3	-$536*** ($90)

Source: Medicare claims and enrollment data for 2010–2022 obtained from the Chronic Conditions Warehouse for IAH and matched comparison group beneficiaries in all IAH practices included in the evaluation.

Notes: Standard errors are in parentheses. We computed coefficients and standard errors by using the weighted sample size, which considers the matching and eligibility weights. Because of rounding, a difference-in-differences estimate displayed as zero might be shown alongside a percentage effect that exceeds zero.

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

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

ADLs = activities of daily living; SE = standard error.

Exhibit D.20. Estimated effect on total Medicare spending by dual eligibility, one, two, and three years after starting HBPC

Period	Non-dually eligible for Medicare and Medicaid				Dually eligible for Medicare and Medicaid				Difference in estimated effects, subgroup vs. reference (SE)
	IAH	Comparison	Difference-in-differences estimated effect (SE)	Percentage effect[a]	IAH	Comparison	Difference-in-differences estimated effect (SE)	Percentage effect[a]	
Third year	$2,916	$2,748	$5 ($204)	0.1	$3,577	$3,503	-$260 ($262)	-4.4	-$265 ($199)
Second year	$3,085	$2,874	$48 ($160)	0.9	$3,825	$3,614	-$123 ($238)	-2.1	-$171 ($235)
First year	$3,865	$4,113	-$411*** ($128)	-8.0	$4,520	$4,959	-$772*** ($229)	-13.2	-$361 ($231)
One year pre-index	$5,128	$4,966	-	-	$5,865	$5,531	-	-	-
Two years pre-index	$1,850	$1,763	-$76 ($167)	-1.5	$2,687	$2,530	-$177 ($198)	-3.0	-$101 ($128)

Source: Medicare claims and enrollment data for 2010–2022 obtained from the Chronic Conditions Warehouse for IAH and matched comparison group beneficiaries in all IAH practices included in the evaluation.

Notes: Standard errors are in parentheses. We computed coefficients and standard errors by using the weighted sample size, which considers the matching and eligibility weights. Because of rounding, a difference-in-differences estimate displayed as zero might be shown alongside a percentage effect that exceeds zero.

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

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

SE = standard error.

Exhibit D.21. Estimated effect on total Medicare spending by combination of dual eligibility for Medicare and Medicaid and functional impairment, one, two, and three years after starting HBPC

Period	Dually eligible for Medicare and Medicaid and needed help with 5–6 ADLs		Non-dually eligible for Medicare and Medicaid and needed help with 5–6 ADLs		Dually eligible for Medicare and Medicaid and needed help with 2–4 ADLs		Non-dually eligible for Medicare and Medicaid and needed help with 2–4 ADLs	
	Differences-in-differences estimated effect (SE)	Percentage effect	Differences-in-differences estimated effect (SE)	Percentage effect[a]	Differences-in-differences estimated effect (SE)	Percentage effect[a]	Differences-in-differences estimated effect (SE)	Percentage effect
Third year	-$707*** ($258)	-11.1	-$387* ($224)	-7.1	$244 ($334)	4.7	$515*** ($180)	11.1
Second year	-$496* ($289)	-7.8	-$143 ($178)	-2.6	$322 ($326)	6.2	$339** ($145)	7.3
First year	-$1,371*** ($232)	-21.5	-$927*** ($149)	-17.0	-$143 ($290)	-2.8	$155 ($191)	3.3
One year pre-index	-	-	-	-	-	-	-	-
Two years pre-index	-$435* ($229)	-6.8	-$256 ($172)	-4.7	$224 ($224)	4.3	$227 ($169)	4.9

Source: Medicare claims and enrollment data for 2010–2022 obtained from the Chronic Conditions Warehouse for IAH and matched comparison group beneficiaries in all IAH practices included in the evaluation.

Notes: Standard errors are in parentheses. We computed coefficients and standard errors by using the weighted sample size, which considers the matching and eligibility weights. Because of rounding, a difference-in-differences estimate displayed as zero might be shown alongside a percentage effect that exceeds zero.

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

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

ADLs = activities of daily living; SE = standard error.

Exhibit D.22. Estimated effect on total Medicare spending and inpatient spending, excluding one practice at a time, one, two, and three years after starting HBPC

Excluding practice	First year		Second year		Third year	
	Difference-in-differences estimated effect (SE)	Percentage effect[a]	Difference-in-differences estimated effect (SE)	Percentage effect[a]	Difference-in-differences estimated effect (SE)	Percentage effect
Total Medicare spending ($, PBPM)						
Full sample	-$501*** ($127)	-9.4	$7 ($151)	0.1	-$68 ($204)	-1.3
A	-$532*** ($137)	-10.1	-$9 ($164)	-0.2	-$81 ($226)	-1.5
B	-$477*** ($145)	-9.0	$16 ($163)	0.3	-$46 ($219)	-0.9
C	-$647*** ($82)	-11.3	-$83 ($100)	-1.5	-$218** ($105)	-3.8
D	-$505*** ($132)	-9.7	-$21 ($157)	-0.4	-$112 ($220)	-2.1
E	-$498*** ($127)	-9.3	$20 ($151)	0.4	-$77 ($207)	-1.5
F	-$502*** ($132)	-9.4	$20 ($156)	0.4	-$68 ($211)	-1.3
G	-$476*** ($136)	-8.8	$36 ($162)	0.7	-$23 ($216)	-0.4
H	-$493*** ($127)	-9.2	$13 ($151)	0.2	-$63 ($206)	-1.2
I	-$464*** ($128)	-8.7	$37 ($147)	0.7	-$36 ($201)	-0.7
J	-$480*** ($128)	-8.9	$18 ($154)	0.3	-$49 ($208)	-0.9
K	-$486*** ($127)	-9.1	$26 ($152)	0.5	-$55 ($207)	-1.0
L	-$514*** ($133)	-9.6	-$10 ($158)	-0.2	-$74 ($213)	-1.4
M	-$455*** ($119)	-8.6	$10 ($149)	0.2	-$41 ($201)	-0.8
N	-$507*** ($135)	-9.5	$14 ($160)	0.3	-$52 ($216)	-1.0
O	-$509*** ($139)	-9.5	-$8 ($164)	-0.1	-$67 ($218)	-1.3
P	-$501*** ($130)	-9.4	$20 ($155)	0.4	-$58 ($213)	-1.1
Inpatient spending ($, PBPM)						
Full sample	-$517*** ($78)	-20.5	-$158 ($97)	-6.3	-$179 ($115)	-7.1
A	-$518*** ($77)	-21.0	-$147 ($103)	-6.0	-$171 ($124)	-6.9
B	-$486*** ($82)	-19.6	-$136 ($94)	-5.5	-$153 ($115)	-6.1
C	-$588*** ($59)	-21.2	-$195*** ($63)	-7.0	-$240*** ($65)	-8.7
D	-$526*** ($85)	-21.5	-$193* ($107)	-7.9	-$233* ($127)	-9.5
E	-$515*** ($78)	-20.5	-$149 ($97)	-5.9	-$186 ($117)	-7.4
F	-$526*** ($80)	-20.8	-$154 ($100)	-6.1	-$182 ($119)	-7.2
G	-$511*** ($84)	-20.0	-$141 ($104)	-5.5	-$164 ($123)	-6.4
H	-$513*** ($78)	-20.3	-$155 ($98)	-6.1	-$176 ($115)	-7.0
I	-$496*** ($79)	-19.8	-$147 ($97)	-5.9	-$158 ($113)	-6.3
J	-$512*** ($80)	-20.1	-$159 ($100)	-6.2	-$178 ($118)	-7.0
K	-$513*** ($79)	-20.4	-$147 ($98)	-5.8	-$175 ($116)	-6.9
L	-$528*** ($81)	-21.0	-$173* ($101)	-6.8	-$179 ($118)	-7.1
M	-$496*** ($78)	-19.9	-$174* ($101)	-7.0	-$166 ($115)	-6.7

Exhibit D.22 (*continued*)

Excluding practice	First year		Second year		Third year	
	Difference-in-differences estimated effect (SE)	Percentage effect[a]	Difference-in-differences estimated effect (SE)	Percentage effect[a]	Difference-in-differences estimated effect (SE)	Percentage effect
N	-$514*** ($83)	-20.4	-$149 ($103)	-5.9	-$164 ($121)	-6.5
O	-$523*** ($85)	-20.8	-$167 ($105)	-6.6	-$181 ($124)	-7.2
P	-$513*** ($79)	-20.6	-$143 ($98)	-5.7	-$173 ($118)	-7.0

Source: Medicare claims and enrollment data for 2010–2022 obtained from the Chronic Conditions Warehouse for IAH and matched comparison group beneficiaries in all IAH practices included in the evaluation.

Notes: Standard errors are in parentheses. We computed coefficients and standard errors by using the weighted sample size, which considers the matching and eligibility weights. Because of rounding, a difference-in-differences estimate displayed as zero might be shown alongside a percentage effect that exceeds zero.

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

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

PBPM = per beneficiary per month; SE = standard error.

Exhibit D.23. Estimated effect on total Medicare spending with outliers truncated, one, two, and three years after starting HBPC

Period	IAH	Comparison	Difference-in-differences estimated effect (SE)	Main difference-in-differences estimated effect (SE)	Percentage effect[a]
Third year	$3,021	$2,872	$9 ($184)	-$68 ($204)	0.2
Second year	$3,183	$2,989	$54 ($144)	$7 ($151)	1.0
First year	$3,937	$4,188	-$392*** ($124)	-$501*** ($127)	-7.5
One year pre-index	$5,218	$5,078	-	-	-
Two years pre-index	$2,074	$1,983	-$49 ($146)	-$99 ($164)	-0.9

Source: Medicare claims and enrollment data for 2010–2022 obtained from the Chronic Conditions Warehouse for IAH and matched comparison group beneficiaries in all IAH practices included in the evaluation.

Notes: Standard errors are in parentheses. We computed coefficients and standard errors by using the weighted sample size, which considers the matching and eligibility weights. Because of rounding, a difference-in-differences estimate displayed as zero might be shown alongside a percentage effect that exceeds zero. We winsorized outliers at the 99th percentile, as described in Section 4.2.7.

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

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

PBPM = per beneficiary per month; SE = standard error.

Exhibit D.24. Estimated effect on spending categories, one, two, and three years after starting HBPC

Period	IAH	Comparison	Difference (IAH – comparison)	Difference-in-differences estimated effect (SE)	Percentage effect[a]
Inpatient spending ($, PBPM)					
Third year	$1,151	$1,201	-$50	-$179 ($115)	-7.1
Second year	$1,256	$1,284	-$28	-$158 ($97)	-6.3
First year	$1,644	$2,032	-$387	-$517*** ($78)	-20.5
One year pre-index	$2,522	$2,393	$130	-	-
Two years pre-index	$825	$761	$64	-$66 ($135)	-2.6
Skilled nursing facility spending ($, PBPM)					
Third year	$405	$458	-$53	-$184*** ($41)	-15.8
Second year	$434	$484	-$50	-$181*** ($38)	-15.6
First year	$543	$747	-$204	-$335*** ($36)	-28.8
One year pre-index	$1,162	$1,031	$130	-	-
Two years pre-index	$253	$167	$86	-$44 ($38)	-3.8
Home health spending ($, PBPM)					
Third year	$320	$238	$81	$69*** ($16)	13.6
Second year	$352	$264	$88	$75*** ($16)	14.8
First year	$556	$374	$183	$170*** ($28)	33.6
One year pre-index	$507	$494	$13	-	-
Two years pre-index	$246	$219	$27	$14 ($11)	2.8
Hospice spending ($, PBPM)					
Third year	$451	$232	$219	$195*** ($47)	373.3
Second year	$423	$196	$227	$203*** ($34)	389.1
First year	$309	$135	$174	$150*** ($27)	287.5
One year pre-index	$52	$29	$24	-	-
Two years pre-index	$33	$37	-$4	-$28*** ($7)	-53.3
Clinician/supplier spending ($, PBPM)					
Third year	$336	$370	-$34	$25 ($28)	4.6
Second year	$375	$384	-$9	$50** ($25)	9.2
First year	$462	$520	-$58	$1 ($26)	0.1
One year pre-index	$542	$601	-$59	-	-
Two years pre-index	$306	$349	-$44	$15 ($19)	2.8

Exhibit D.24 (*continued*)

Period	IAH	Comparison	Difference (IAH – comparison)	Difference-in-differences estimated effect (SE)	Percentage effect[a]
Durable medical equipment spending ($, PBPM)					
Third year	$49	$47	$1	$11** ($5)	14.6
Second year	$52	$49	$2	$12** ($5)	15.9
First year	$74	$62	$13	$22*** ($5)	30.4
One year pre-index	$73	$82	-$9	-	-
Two years pre-index	$63	$66	-$4	$5** ($2)	7.4
Outpatient spending ($, PBPM)					
Third year	$225	$295	-$70	-$24 ($19)	-7.7
Second year	$238	$302	-$64	-$19 ($16)	-6.0
First year	$249	$342	-$93	-$47** ($23)	-15.2
One year pre-index	$313	$358	-$45	-	-
Two years pre-index	$262	$287	-$25	$21*** ($7)	6.6

Source: Medicare claims and enrollment data for 2010–2022 obtained from the Chronic Conditions Warehouse for IAH and matched comparison group beneficiaries in all IAH practices included in the evaluation.

Notes: Standard errors are in parentheses. We computed coefficients and standard errors by using the weighted sample size, which considers the matching and eligibility weights. Because of rounding, a difference-in-differences estimate displayed as zero might be shown alongside a percentage effect that exceeds zero.

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

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

PBPM = per beneficiary per month; SE = standard error.

Exhibit D.25. Estimated effect on inpatient spending by functional impairment (ADLs), one, two, and three years after starting HBPC

Period	Needed help with 2–4 ADLs				Needed help with 5–6 ADLs				Difference in estimated effects, subgroup vs. reference (SE)
	IAH	Comparison	Difference-in-differences estimated effect (SE)	Percentage effect[a]	IAH	Comparison	Difference-in-differences estimated effect (SE)	Percentage effect[a]	
Third year	$1,105	$1,081	$126 ($104)	5.6	$1,129	$1,252	-$420*** ($130)	-15.5	-$546*** ($98)
Second year	$1,078	$1,111	$70 ($89)	3.1	$1,369	$1,392	-$319** ($126)	-11.8	-$389*** ($119)
First year	$1,031	$1,287	-$153 ($117)	-6.8	$2,062	$2,604	-$838*** ($84)	-31.0	-$685*** ($109)
One year pre-index	$2,241	$2,345	-	-	$2,705	$2,409	-	-	-
Two years pre-index	$724	$663	$164 ($136)	7.3	$899	$811	-$208 ($137)	-7.7	-$372*** ($70)

Source: Medicare claims and enrollment data for 2010–2022 obtained from the Chronic Conditions Warehouse for IAH and matched comparison group beneficiaries in all IAH practices included in the evaluation.

Notes: Standard errors are in parentheses. We computed coefficients and standard errors by using the weighted sample size, which considers the matching and eligibility weights. Because of rounding, a difference-in-differences estimate displayed as zero might be shown alongside a percentage effect that exceeds zero.

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

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

ADLs = activities of daily living; SE = standard error.

Exhibit D.26. Estimated effect on inpatient spending by dual eligibility for Medicare and Medicaid, one, two, and three years after starting HBPC

Period	Non-dually for Medicare and Medicaid eligible				Dually eligible for Medicare and Medicaid				Difference in estimated effects, subgroup vs. reference (SE)
	IAH	Comparison	Difference-in-differences estimated effect (SE)	Percentage effect[a]	IAH	Comparison	Difference-in-differences estimated effect (SE)	Percentage effect[a]	
Third year	$990	$1,055	-$150 ($111)	-6.5	$1,589	$1,586	-$277 ($169)	-9.2	-$127 ($140)
Second year	$1,074	$1,143	-$153 ($100)	-6.6	$1,729	$1,652	-$204 ($160)	-6.8	-$51 ($155)
First year	$1,493	$1,856	-$448*** ($75)	-19.2	$2,047	$2,500	-$734*** ($167)	-24.3	-$286 ($175)
One year pre-index	$2,328	$2,243	-	-	$3,021	$2,741	-	-	-
Two years pre-index	$677	$632	-$40 ($130)	-1.7	$1,169	$1,026	-$137 ($156)	-4.5	-$97 ($93)

Source: Medicare claims and enrollment data for 2010–2022 obtained from the Chronic Conditions Warehouse for IAH and matched comparison group beneficiaries in all IAH practices included in the evaluation.

Notes: Standard errors are in parentheses. We computed coefficients and standard errors by using the weighted sample size, which considers the matching and eligibility weights. Because of rounding, a difference-in-differences estimate displayed as zero might be shown alongside a percentage effect that exceeds zero.

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

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

SE = standard error.

Exhibit D.27. Estimated effect on inpatient spending by combination of dual eligibility for Medicare and Medicaid and functional impairment (ADLs), one, two, and three years after starting HBPC

Period	Dually eligible for Medicare and Medicaid and needed help with 5–6 ADLs		Non-dually eligible for Medicare and Medicaid and needed help with 5–6 ADLs		Dually eligible for Medicare and Medicaid and needed help with 2–4 ADLs		Non-dually eligible for Medicare and Medicaid and needed help with 2–4 ADLs	
	Differences-in-differences estimated effect (SE)	Percentage effect[a]	Differences-in-differences estimated effect (SE)	Percentage effect[a]	Differences-in-differences estimated effect (SE)	Percentage effect[a]	Differences-in-differences estimated effect (SE)	Percentage effect[a]
Third year	-$584*** ($199)	-17.4	-$376*** ($129)	-15.2	$87 ($183)	3.4	$137 ($95)	6.5
Second year	-$446* ($230)	-13.3	-$292** ($118)	-11.8	$101 ($186)	3.9	$44 ($91)	2.1
First year	-$1,158*** ($185)	-34.5	-$746*** ($97)	-30.2	-$266 ($193)	-10.4	-$113 ($109)	-5.4
One year pre-index	-	-	-	-	-	-	-	-
Two years pre-index	-$400** ($185)	-11.9	-$145 ($128)	-5.9	$254 ($162)	9.9	$131 ($139)	6.2

Source: Medicare claims and enrollment data for 2010–2022 obtained from the Chronic Conditions Warehouse for IAH and matched comparison group beneficiaries in all IAH practices included in the evaluation.

Notes: Standard errors are in parentheses. We computed coefficients and standard errors by using the weighted sample size, which considers the matching and eligibility weights. Because of rounding, a difference-in-differences estimate displayed as zero might be shown alongside a percentage effect that exceeds zero.

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

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

ADLs = activities of daily living; SE = standard error.

Exhibit D.28. Hospice use among IAH and comparison beneficiaries who did and did not die in each year, percentage

Year	Remained alive until the end of the year		Died in the year	
	IAH	Comparison	IAH	Comparison
Third year	11.3	5.8	65.2	56.9
Second year	11.4	5.1	64.2	54.4
First year	10.1	4.1	69.1	57.1

Source: Medicare claims and enrollment data for 2010–2022 obtained from the Chronic Conditions Warehouse for IAH and matched comparison group beneficiaries in all IAH practices included in the evaluation.

Note: Results are not regression-adjusted.

IAH = Independence at Home; SE = standard error.

Exhibit D.29. Estimated effect on days at home, one, two, and three years after starting HBPC (percentage point change)

Period	IAH	Comparison	Difference (IAH – comparison)	Difference-in-differences estimated effect (SE)	Percentage effect[a]
Third year	93.06	92.58	0.48	2.47*** (0.33)	2.9
Second year	93.09	92.37	0.72	2.71*** (0.28)	3.2
First year	92.21	89.93	2.27	4.26*** (0.30)	5.0
One year pre-index	85.21	87.20	-1.99	-	-
Two years pre-index	96.07	96.96	-0.88	1.11*** (0.23)	1.3

Source: Medicare claims and enrollment data for 2010–2022 obtained from the Chronic Conditions Warehouse for IAH and matched comparison group beneficiaries in all IAH practices included in the evaluation.

Notes: Standard errors are in parentheses. We computed coefficients and standard errors by using the weighted sample size, which considers the matching and eligibility weights. Because of rounding, a difference-in-differences estimate displayed as zero might be shown alongside a percentage effect that exceeds zero.

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

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

SE = standard error.

Exhibit D.30. Estimated effect on the probability of nursing home entry, one, two, and three years after starting HBPC (percentage point change)

Period	IAH	Comparison	Difference (IAH – comparison)	Difference-in-differences estimated effect (SE)	Percentage effect[a]
Third year	16.03	14.10	1.93	-5.07** (2.25)	-33.2
Second year	11.69	11.78	-0.08	-7.09*** (1.38)	-46.4
First year	6.85	8.47	-1.62	-8.63*** (1.01)	-56.4
One year pre-index	15.29	8.28	7.01	-	-
Two years pre-index	3.11	1.39	1.72	-5.29*** (0.70)	-34.6

Source: Medicare claims and enrollment data for 2010–2022 obtained from the Chronic Conditions Warehouse for IAH and matched comparison group beneficiaries in all IAH practices included in the evaluation.

Notes: Standard errors are in parentheses. We computed coefficients and standard errors by using the weighted sample size, which considers the matching and eligibility weights. Because of rounding, a difference-in-differences estimate displayed as zero might be shown alongside a percentage effect that exceeds zero.

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

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

SE = standard error.

Exhibit D.31. Estimated association with the probability of dying, one, two, and three years after starting HBPC (percentage point change)

Period	IAH	Comparison	Difference IAH – comparison (SE)
Third year	68.73	62.44	6.29*** (0.34)
Second year	51.33	45.67	5.66*** (0.37)
First year	28.21	24.73	3.48*** (0.28)

Source: Medicare claims and enrollment data for 2010–2022 obtained from the Chronic Conditions Warehouse for IAH and matched comparison group beneficiaries in all IAH practices included in the evaluation.

Note: Results from this study do *not* suggest that IAH inappropriately caused beneficiaries to die more quickly. For more information, please see the report which describes several factors that could have contributed to the results shown in this exhibit. Standard errors are in parentheses. We computed means, differences, and standard errors by using the weighted sample size, which considers the matching and eligibility weights.

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

SE = standard error.

Exhibit D.32. Results for the performance-year analysis and three-year analysis

Outcome	Performance-year analysis	Three-year analysis		
Primary care visits	7.4% (Years 1–10)	154.1%* (First year)	119.9%* (Second year)	104.7%* (Third year)
Specialty care visits	-9.3% (Years 1–10)	-9.4%* (First year)	4.4% (Second year)	7.9%* (Third year)
Advanced imaging spending	-1.1% (Years 1–10)	-10.2%* (First year)	-1.6% (Second year)	-2.0% (Third year)
Testing spending	-2.9% (Years 1–10)	61.7%* (First year)	40.2%* (Second year)	39.1%* (Third year)
Other imaging spending	-10.0%* (Years 1–10)	26.9%* (First year)	23.8%* (Second year)	19.8%* (Third year)
Chronic care management services	n.a.	22pp* (First year)	18pp* (Second year)	19pp* (Third year)
Outpatient ED visits	-1.7% (Years 1–10)	-6.5% (First year)	-9.3%* (Second year)	-12.1%* (Third year)
Potentially avoidable ED visits	1.7% (Years 1–10)	-13.8%* (First year)	-16.0%* (Second year)	-16.3%* (Third year)
Hospital admissions preceded by an ED visit	-5.8%* (Years 1–10)	-10.9%* (First year)	-5.6%* (Second year)	-4.5%* (Third year)
Hospital admissions	-2.3% (Years 1–10)	-11.4%* (First year)	-5.0%* (Second year)	-3.2% (Third year)
Potentially avoidable hospital admissions	-3.9% (Years 1–10)	-11.1%* (First year)	-8.5%* (Second year)	-3.7% (Third year)
Unplanned readmissions	-4.3% (Years 1–10)	-19.8%* (First year)	-7.2%* (Second year)	-7.9%* (Third year)
Total Medicare spending before incentive payments	-4.9% (Years 1–10)	-9.4%* (First year)	0.1% (Second year)	-1.3% (Third year)
Days at home	0.1% (Years 1–10)	5.0%* (First year)	3.2%* (Second year)	2.9%* (Third year)
Nursing home entry	9.3%* (Years 1–10)	-56.4%* (First year)	-46.4%* (Second year)	-33.2%* (Third year)
Death	-3.8% (Years 1–10)	3.5pp* (First year)	5.7pp* (Second year)	6.3pp* (Third year)

Exhibit D.32 (*continued*)

Note: This exhibit presents results for the performance-year analysis and three-year analysis. For each outcome, the exhibit reports the impact as a percentage change (or percentage point change for two outcomes). It uses symbols to denote the strength of the evidence for each outcome. The four levels of evidence are increased/decreased, may have increased/decreased, probably did not increase/decrease, and did not increase/decrease. The exhibit also uses asterisks to indicate if an impact estimate is statistically significant ($p < 0.10$). All results are statistically significant when the level of evidence indicates that there was an increase/decrease in the outcome. Sometimes results are statistically significant when the level of evidence indicates that there may have been an increase/decrease in the outcome or there probably was not an increase/decrease in the outcome. The results from the performance year analysis report the estimated impacts of the demonstration. The results from the three-year analysis report the estimated impacts of HBPC during the demonstration.

pp = percentage point change.

References

Pimentel, S.D., L.V. Forrow, J. Gellar, and J. Li. "Optimal Matching Approaches in Health Policy Evaluations Under Rolling Enrollment." *Journal of the Royal Statistical Society: Series A (Statistics in Society)*, vol. 183, no. 4, 2020, pp. 1411–1435.

Kim, D.H., S. Schneeweiss, R.J. Glynn, L.A. Lipsitz, K. Rockwood, and J. Avorn. "Measuring Frailty in Medicare Data: Development and Validation of a Claims-Based Frailty Index." *The Journals of Gerontology Series A: Biological Sciences and Medical Sciences*, vol. 73, no. 7, 2018, pp. 980–987.

Kim, D.H., R.J. Glynn, J. Avorn, L.A. Lipsitz, K. Rockwood, A. Pawar, and S. Schneeweiss. "Validation of a Claims-Based Frailty Index Against Physical Performance and Adverse Health Outcomes in the Health and Retirement Study." *The Journals of Gerontology Series A: Biological Sciences and Medical Sciences*, vol. 74, no. 8, 2019, pp. 1271–1276.

Kim, D.H., E. Patorno, A. Pawar, H. Lee, S. Schneeweiss, and R.J. Glynn. "Measuring Frailty in Administrative Claims Data: Comparative Performance of Four Claims-Based Frailty Measures in the U.S. Medicare Data." *The Journals of Gerontology Series A: Biological Sciences and Medical Sciences*, vol. 75, no. 6, 2020, pp. 1120–1125.

Kim, D.H., and N. Gautam. "SAS Programs - Claims-Based Frailty Index." Harvard Dataverse. 2020. https://doi.org/doi:10.7910/DVN/HM8DOI.

Kim, D.H., and N. Gautam. "Research Background - Claims-Based Frailty Index 20240614.docx [online]." V6 ed. SAS Programs - Claims-Based Frailty Index Harvard Dataverse; 2024. https://doi.org/10.7910/DVN/HM8DOI/HVLESQ.

Appendix E.

Additional Methods, Supplementary Results, and Limitations

1. Overview

This appendix describes additional methods, supplementary results, and limitations of the Independence at Home (IAH) evaluation. First, we describe the methods used to collect and analyze survey to understand the experiences of beneficiaries and caregivers. Second, we describe the methods used to collect and analyze qualitative data to understand how IAH practices changed home-based primary care (HBPC) delivery over time. Third, we describe the methods we used to analyze and identify best practices for reducing potentially avoidable hospital admissions. Third, we include tables of supplementary results related to quality measure performance, level of IAH incentive payments earned, and best practices. Finally, we describe limitations of the IAH evaluation.

2. Survey data and methods

To answer research questions about beneficiaries' and caregivers' experiences, we conducted two interconnected, self-administered surveys distributed by mail: (1) a survey of beneficiaries currently or previously enrolled in the IAH demonstration (or their proxies), excluding beneficiaries who were deceased or who had moved from the area, and (2) a survey of the beneficiaries' caregivers.

Together, the surveys measured the following domains: experiences of care, satisfaction, clinician attributes, self-reported health outcomes, beneficiary characteristics, and caregiver attributes. After identifying the key domains and subdomains, we reviewed existing questionnaires to identify survey items measuring the domains of interest. We prioritized survey items from the Medicare Current Beneficiary Survey, because the questions were designed for Medicare beneficiaries. If that survey lacked adequate coverage for a particular domain or subdomain, we adapted items from the questionnaires used in the Consumer Assessment of Healthcare Providers and Systems and the Medicare Care Management Performance Demonstration evaluation (Dale et al. 2012), which Mathematica designed and collected for CMS.

For questions specific to the caregiver survey, we adopted or adapted items from the questionnaires used in (1) Mathematica's evaluation of the Cash and Counseling demonstration (Brown et al. 2007) and (2) a study of unmet needs of adults with disabilities conducted by the Massachusetts Department of Public Health and the Center for Survey Research at the University of Massachusetts–Boston (2008). We overlapped the content covered by the beneficiary and caregiver questions, because they provided different perspectives on experience in the IAH program. In some cases, we developed our own questions specific to the evaluation, such as whether beneficiaries would contact their clinicians if they did not feel well and were unsure whether they should go to the ED.

We created a 15-minute questionnaire for beneficiaries and a 10-minute questionnaire for caregivers. After developing these surveys, we conducted a small, iterative, two-phase pretest to evaluate whether respondents had difficulty completing the questionnaires. We translated all survey materials into Spanish and made the translated materials available to beneficiaries and their caregivers upon request.

The original size of the beneficiary sample was 7,293 beneficiaries whose claims-based enrollment date, as reported by the implementation and monitoring contractor, occurred from June 1, 2012, to June 30, 2014. The implementation and monitoring contractor determined the claims-based enrollment date as the date on which the beneficiary first met the Medicare enrollment, hospitalization, and rehabilitation services criteria for the demonstration. All of these beneficiaries were enrolled and eligible according to the IAH practice. Beneficiaries who died before the survey was fielded did not receive a survey, though we used them to estimate the number of dead people among the unlocated or nonrespondent sample members.

We conducted the survey in seven cohorts during IAH Years 1 to 3. Cohort 1 included beneficiaries whose claims-based enrollment date was June 1, 2012, to December 31, 2012. Cohort 2 included beneficiaries whose claims-based enrollment date was January 1, 2013, to March 31, 2013. Similarly, all other cohorts consisted of beneficiaries enrolled in a three-month period ending with Cohort 7, which included beneficiaries whose claims-based enrollment date was April 1, 2014, to June 30, 2014.

The initial mailing included a cover letter, the beneficiary and caregiver surveys, a survey fact sheet, and two postage-paid return envelopes. If the respondent did not return one or both surveys, we attempted a series of subsequent contacts, including three postcards and two additional survey packets. All beneficiaries who did not respond to the mailings were selected for telephone follow-up. Interviewers attempted to complete the survey on the telephone with the beneficiary. We mailed an additional paper survey to the caregiver of each beneficiary who completed the survey by telephone.

The survey data presented in the report reflect responses from 3,870 beneficiaries (a response rate of 63.3 percent) and 2,519 caregivers (a response rate of 55.7 percent), and we present responses weighted for nonresponse. We calculated the response rate after excluding beneficiaries who could not respond, such as those who had a language barrier or who passed away before or during data collection. We calculated separate nonresponse weights for beneficiaries and caregivers. For each group, we calculated the weights in three stages: (1) eligibility determination, (2) location adjustment, and (3) cooperation adjustment. We used the eligibility model to estimate the probability of survival among unlocated cases. We calculated the location and cooperation adjustments using weighting classes based upon propensity scores from logistic models. The location adjustment accounts for people who were sampled but could not be located during data collection. The weights of these people were reallocated to similar sample members who were located (both respondents and nonrespondents). Their eligibility was assumed to be unknown.

3. Qualitative data collection and methods

When reporting information about the IAH practices in the report and appendices, we relied on qualitative data gathered from demonstration practices during demonstration Years 1 to 3, Year 5, and Year 6. We collected these data to understand why and how the opportunity to earn performance-based incentive payments may have affected outcomes. Specifically, qualitative interviews sought to understand how IAH practices' delivery of HBPC changed throughout the demonstration, which required identifying how practices provided HBPC before and after joining the demonstration. Below we describe the rounds of qualitative data collection, ordered from most recent to least.

- We conducted the final round of interviews from November 2019 to February 2020, spanning the transition from Year 6 to Year 7. 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 HarmonyCares Medical Group (known at that time as Visiting Physicians Association) to collect additional insights on centralized supports for these 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 previous round of interviews in April 2017 (during Year 5) and fall 2019 (during Year 6), including their efforts associated with meeting performance thresholds for IAH quality measures. We also asked about motivations for any changes and factors that affected implementation of those changes. These interviews took place before the COVID-19 pandemic.
- In April 2017, we conducted telephone interviews with 25 clinical and administrative staff at 15 IAH practices and the HarmonyCares Medical Group 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 emergency department (ED) use, coordinate care, ensure round-the-clock access to care, follow up with beneficiaries and reconcile medications within 48 hours after discharge from the hospital or ED, and document beneficiaries' preferences. We also asked about their 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 the practices delivered 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. Care partners were organizations external to the IAH practice's care team with which the practice had an established working relationship to coordinate care for beneficiaries. We asked respondents questions that enabled us to assess how care partner organizations perceived IAH practices on three key dimensions of HBPC: care coordination, accessibility, and continuity.

- During visits to demonstration practices from April 2015 to October 2015 (late in Year 3 and early in Year 4), we interviewed the practices' 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 practices planned to sustain the HBPC model.
- We conducted two rounds of early site visits from February to May 2013 (Year 1) and February to July 2014 (Year 2). During these early 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. We also collected information on barriers to and facilitators of meeting the requirements of the demonstration and on how practices used information technologies, such as electronic health records and health information exchanges 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 beneficiaries 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 practices, 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 demonstration; and (5) the processes used to implement the demonstration. 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 practice and identified common themes across practices.

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 practices until after the demonstration began, data on practices' operations and structure before the demonstration were limited to what interviewees told us was different in Year 1 relative to before the demonstration.

4. Methods for analysis of best practices under the demonstration

The IAH legislation states that "the Secretary shall conduct an independent evaluation of the demonstration and submit to Congress a final report, including best practices under the demonstration, including, to the extent practicable, with respect to the use of electronic health information systems." (See Appendix A for the legislation.)

To address this aspect of the legislation, we conducted a best practices analysis to assess the relationship between a range of care delivery activities and potentially avoidable hospital admissions. We focused on potentially avoidable hospital admissions because these admissions incur high costs for Medicare and can sometimes be avoided through high-quality, timely primary care. Additionally, hospital admissions may negatively affect older adults, in some cases leading to functional decline, adverse drug reactions, delirium, and sometimes even permanent disability and nursing home placement.

This section describes the mixed-methods approach to assessing the relationships between care delivery activities and potentially avoidable hospital admissions. First, we describe how we estimated the effect of starting HBPC during the demonstration on potentially avoidable hospital admissions for each IAH practice, as measured in the three-year analysis (see Appendix D for more details). For each care delivery activity, we then describe how we compared outcomes between groups to identify which care delivery activities had statistically significant effects. Finally, we describe how we explored structural characteristics associated with high-performing practices to understand contextual factors that may support effective care delivery. Together, these mixed methods produced the best practices analysis.

Estimating the effect for each IAH practice. We began by estimating the effect of starting HBPC during the demonstration on potentially avoidable hospital admissions for each of the 16 IAH practices by pooling estimated effects across all the performance years that the practice participated in the IAH demonstration, resulting in one estimate per practice.[46] For example, because the practice in Cleveland left the demonstration after IAH Year 5, we estimated the average annual effect across five years. Each practice effect was estimated using the same IAH and comparison groups and the same regression approach as the three-year analysis (see Appendix D for details). To calculate total beneficiary-months for each practice, we counted the total number of months that each beneficiary was eligible for IAH during the time period that the practice participated in IAH.

Categorizing practices by care delivery activity. To be considered a best practice, we needed to hypothesize that the care delivery activity could be associated with fewer potentially avoidable hospital admissions. For example, we hypothesized that providing after-hours and weekend visits may be associated with fewer potentially avoidable hospital admissions, because beneficiaries could receive urgent care from the IAH practice at home after hours.

[46] We calculated estimates separately for each of the three members of the Mid-Atlantic Consortium (located in Richmond, Viginia; Washington, DC/Baltimore, Maryland; and Philadelphia, Pennsylvania), because they had some differences in care delivery activities and structural characteristics that we examined in the best practices analysis. For more information, see Appendix B.

We considered a range of care delivery activities for this analysis. Information about practices' care delivery activities came from practice-reported data about their care delivery activities either submitted to the IAH implementation contractor or Mathematica, descriptive analyses of claims data, as well as qualitative data from structured interviews conducted in Years 1, 5, and 6.

For each care delivery activity, we classified practices into Group A or Group B according to differentiating criteria unique to each care delivery activity. Exhibit E.1 lists the care delivery activities assessed for the best practices analysis, as well as the criteria needed to be classified into Group A or B. When analyzing data to assign practices to groups, we considered the entire demonstration period, meaning practices did not change groups in different performance years. In some cases, a practice's care delivery activity changed between performance years. In those cases, we assigned the practice a group based on an average of its overall performance, such as number of years it met criteria for Group A versus Group B and how close the practice was to the threshold for Group A versus B for care delivery activities with nonbinary criteria.

Exhibit E.1. Criteria used to group practices and number of practices per group for each care delivery activity studied in the best practices analysis

Care delivery activity	Hypothesis	Data source	Group A[a]		Group B	
			Criteria	N	Criteria	N
Beneficiary visits in settings other than the home[b]	Conducting beneficiary visits in settings other than home, such as hospitals and SNFs, improves clinicians' awareness of their beneficiaries' health status and needs for transitional care after discharge. Improved transitional care may reduce hospital admissions.	Survey	Yes	5	No	11
Daily clinician caseload	Lower daily clinician caseloads may reduce potentially avoidable hospital admissions by allowing clinicians more time per beneficiary.	Survey	3 to 7 beneficiaries	8	8 to 15 beneficiaries	8
Clinician panel size (all beneficiaries, not only those eligible for IAH)	Smaller clinician panel sizes enable closer monitoring, stronger relationships, and more consistent follow-up with each beneficiary. These features may help prevent or identify exacerbations that might otherwise lead to potentially avoidable hospital admissions.	Survey	40 to 149 beneficiaries	6	150 to 200 beneficiaries	10
Used remote monitoring	Using remote monitoring technology allows clinicians to track beneficiaries' health in real time. This may enable timely intervention to prevent exacerbations from progressing to the point of requiring hospital admissions.	Interviews	Yes	6	No or information not available[c]	10
Home visits were conducted by staff other than primary care clinicians, such as registered nurses, social workers, and chaplains	Home visits by staff such as nurses, social workers, and chaplains increase monitoring, reinforce care plans, and address social and emotional needs. This may help identify unmet needs and health risks and prevent conditions from escalating to hospital admissions.	Survey	Yes	12	No	4
Type of primary care clinician conducting home visits	Visits by PAs and NPs may increase access and allow more frequent monitoring and timely interventions, preventing acute issues from escalating into hospital admissions. Also, since many states have laws requiring a physician to supervise or collaborate with PAs and NPs, patients who receive visits from PAs and NPs may benefit from having two clinicians involved with their health care.	Survey	50% or more of visits provided by physicians and less than 50% provided by NPs and PAs	9	50% or more of visits provided by NPs and PAs and less than 50% provided by physicians	7

Exhibit E.1 (*continued*)

Care delivery activity	Hypothesis	Data source	Group A[a]		Group B	
			Criteria	N	Criteria	N
Practice offered weekend visits	Weekend home visits may provide timely care when issues arise outside weekdays, helping manage acute problems early and reducing avoidable hospital admissions.	Survey	Yes	11	No	5
Typical visit frequency for IAH beneficiaries	More frequent home visits may allow earlier detection and management of health issues, improving adherence and preventing conditions from worsening into avoidable hospital admissions.	Survey	High frequency: every 4 weeks	7	Low or variable frequency: every 5 to 12 weeks or no set standard	9
Used formal risk stratification[d]	Risk stratification identifies high-risk beneficiaries for closer monitoring and proactive care. Such identification allows early interventions that may prevent complications and reduce avoidable hospital admissions.	Survey	Yes	8	No (practices may have used clinical judgment to risk stratify beneficiaries)	8
Percentage of primary care visits by telehealth or telephone in the first year of the COVID-19 pandemic	Greater use of telehealth during the first year of the pandemic sustained access to timely primary care, helping manage acute and chronic issues early and reducing avoidable hospital admissions.	Medicare claims	Accounted for 40% or more of visits to IAH beneficiaries	5	Accounted for less than 40% of visits to IAH beneficiaries	7
Mode of hospital admission or ED visit notification	Automated health information exchange notifications give practices timely information on hospital and ED use, enabling rapid follow-up and care coordination that may help prevent avoidable admissions.	Survey	Automated notification through health information exchange for all hospitals/systems in the state	5	Automated notification for *some* hospitals/systems, or practice had no automated notification and relied on beneficiary, caregiver, or assisted living facility	11

Exhibit E.1 (*continued*)

Care delivery activity	Hypothesis	Data source	Group A[a]		Group B	
			Criteria	N	Criteria	N
Performance on practice-reported quality measures used to calculate IAH payment incentive	High performance on these three quality measures means practices were likely to conduct activities that may reduce potentially avoidable hospital admissions, such as follow-up contacts and medication reconciliation within 48 hours of hospital discharges and ED visits/	IAH implementation contractor	Met threshold for at least 70% of the years for at least 2 of the 3 measures	7	Did not meet threshold described in Group A criteria	9
Offered some beneficiaries at least one type of mobile diagnostic technology	Mobile diagnostics such as ultrasounds, x-rays, and blood draws increases access to diagnostic services for beneficiaries who may have limited mobility or transportation. This may enable more timely diagnostics and treatment, allowing early intervention that prevents complications and reduces avoidable hospital admissions.	Interview data	Yes	11	No or information not available[c]	5
Practice offered after-hours visits	After-hours visits increase beneficiary access to timely care when issues arise outside regular office hours. This may help clinicians manage acute problems early, thereby reducing avoidable hospital admissions.	Survey	Yes	6	No	10

Notes: The survey data source refers to a worksheet that IAH practices completed and submitted to Mathematica in Years 5 and 6. The interview data source refers to interviews conducted by Mathematica in Years 1, 5, and 6. The Medicare claims data source includes analysis of Medicare claims data. The IAH implementation contractor data source represents data IAH practices submitted to the Centers for Medicare & Medicaid Services and its implementation contractor.

[a] For all but one care delivery activity, we hypothesized that Group A would be associated with larger reductions in potentially avoidable hospital admissions. The one exception was the type of primary care clinician conducting home visits, where we did not have a hypothesis on what type of clinician would be associated with larger reductions in potentially avoidable hospital admissions.

[b] Examples of beneficiary visits conducted outside the home include those in hospitals and skilled nursing facilities (and, in very limited cases, offices or outpatient clinics).

[c] We identified practices as using remote monitoring or mobile diagnostics based on interview transcripts. Interviewers did not explicitly ask about these technologies; mentions arose organically. If a practice did not mention using the technology, we categorized them in Group B.

[d] In this analysis, we define formal risk stratification as an algorithm-based process to group beneficiaries into high- and low-risk groups to aid in care planning. Some practices risk stratified beneficiaries based on clinician judgment only, which we categorized as informal risk stratification.

ED = emergency department; NP = nurse practitioner; PA = physician assistant.

The care delivery activities included in this analysis were based on the use of remote monitoring and mobile diagnostic technology; structured data we collected in Years 5 and 6; and hypotheses that the activity could be associated with reductions in potentially avoidable hospital admissions. We excluded care delivery areas where practices did not vary in their activities. For example, 15 of the 16 practices reported conducting proactive outreach to some IAH beneficiaries or their caregivers; we did not include this activity in the analysis because there was insufficient variation across the practices.

Estimating effects for Groups A and B for each care delivery activity. For each care delivery activity, we calculated the average estimated effect of starting HBPC during the demonstration on potentially avoidable admissions for Group A and Group B, weighting practice-level results by the number of beneficiary-months. For example, if six of the 16 IAH practices used formal risk stratification, we calculated the weighted average of the effect of starting HBPC during the demonstration among those six practices using beneficiary-months as the weight. This approach produced a single estimated effect of how starting HBPC during the demonstration at an IAH practice that used formal risk stratification affected potentially avoidable hospital admissions. Similarly, we calculated the weighted average among the 10 practices that did not use formal risk stratification.

Comparing estimated effects between groups. For each care delivery activity, we compared the weighted averages of potentially avoidable hospital admissions between Groups A and B to assess whether practices in one group performed substantively better than practices in the other group. We included in that assessment a measure of statistical uncertainty between the groups (confidence intervals and *p*-values). To be considered a best practice, the difference between Group A and B needed to be statistically significant ($p < 0.1$). To determine if the difference between Group A and Group B was statistically significant, we used a single regression model, where an indicator for group was interacted with all other variables (including IAH treatment and year) on the right-hand side of the main regression equation. The triple difference term, which is the interaction between IAH treatment, IAH Year, and group participation, estimated the difference between the effects of starting HBPC during the demonstration in Group A versus Group B, including a test for whether that difference was statistically significant.

We do not assert causality from this analysis because we measured a relatively small number of care delivery activities and lacked data on care delivery activities from non-IAH primary care practices. It is possible, for example, that the practices implementing a particular care delivery activity also shared another unmeasured characteristic, which may have contributed to the observed reduction in potentially avoidable hospital admissions.

Assessing structural characteristics associated with best practices. As an exploratory analysis, we tried to identify contextual features of IAH practices that may have facilitated implementing best practices. First, we identified which practices reported performing at least four best practices. We then qualitatively compared structural characteristics of practices with at least four best practices with the remaining practices. The structural characteristics included profit status, organizational affiliation, participation in an accountable care organization (ACO), whether the practice cared for Medicare Advantage enrollees, and beneficiary residential settings (private residence or group residence, such as an assisted living facility). Details are shown in Exhibit E.5.

5. Supplementary tables

Exhibit E.2. Number and percentage of practices that met the performance threshold for each IAH quality measure by year

	Year 1		Year 2		Year 3		Year 4		Year 5		Year 6		Year 7		Year 8		Year 9		Year 10	
	N	%	N	%	N	%	N	%	N	%	N	%	N	%	N	%	N	%	N	%
Readmissions	14	100	14	100	14	100	14	100	13	93	10	91	10	100	7	100	7	100	1	100
Hospital admission	13	93	13	93	14	100	14	100	14	100	11	100	10	100	7	100	6	86	1	100
ED visits	14	100	12	86	14	100	12	86	13	93	9	82	9	90	7	100	6	86	1	100
Beneficiary preferences	13	93	12	86	11	79	8	57	8	57	3	27	5	50	1	14	2	29	1	100
48-hour contact	5	36	6	43	5	36	6	43	5	36	4	36	4	40	1	14	2	29	1	100
Medication reconciliation	4	29	5	36	5	36	5	36	4	29	4	36	3	30	1	14	1	14	1	100
Total number of practices that participated in IAH	14		14		14		14		14		11		10		7		7		1	

Source: Mathematica's analysis of data from the IAH implementation contractor.

Note: See Exhibit 5.1 in the report for a description of the six measures. To determine performance on these measures, the IAH implementation contractor used Medicare claims and enrollment data, along with site-reported information provided through the IAH Reporting System.

Exhibit E.3. Number and percentage of practices that received full, partial, or no part of the IAH incentive payment by performance year

	Year 1	Year 2	Year 3	Year 4	Year 5	Year 6	Year 7	Year 8	Year 9	Year 10
Practices that participated in IAH	14	14	14	14	14	11	10	7	7	1
Practices that qualified for payment based on spending results	9	7	7	8	8	10	10	7	7	1
Amount of the incentive payment received, among practices that qualified for payment based on spending results										
Full incentive payment	3	4	4	4	3	3	3	1	1	1
Partial incentive payment	6	3	3	3	5	6	7	6	5	0
83%	1	0	0	0	1	0	0	0	0	0
67%	5	3	3	0	1	0	2	0	1	0
50%	0	0	0	3	3	6	5	6	4	0
No incentive payment	0	0	0	1	0	1	0	0	1	0
Percentage that received no payment or partial payment because of quality measure performance	67%	43%	43%	50%	63%	70%	70%	86%	86%	0%

Source: Mathematica's analysis of data from the IAH implementation contractor.

Exhibit E.4. Effects of IAH on potentially avoidable hospital admissions per 1,000 beneficiaries during the first year after beneficiaries started home-based primary care

Care delivery activities	Difference in difference (SE) Group A	Difference in difference (SE) Group B	Triple difference (SE) Group B minus A
Best practices: statistically significant at $p < 0.1$ and effect was in the expected direction			
Beneficiary visits in settings other than the home	-167.6* (35.8)	-39.3* (11.9)	128.3* (37.7)
Daily physician caseload	-122.3* (30.1)	-34.5* (11.9)	87.9* (32.4)
Clinician panel size (all beneficiaries, not only those eligible for IAH)	-111.7* (17.5)	-38.7* (12.5)	73.0* (21.5)
Used remote monitoring	-93.9* (20.1)	-34.4* (12.4)	59.5* (23.6)
Home visits were conducted by staff other than primary care clinicians, such as registered nurses, social workers, and chaplains	-64.9* (15.8)	-15.8* (8.8)	49.1* (18.1)
Type of primary care clinician conducting home visits	-40.4* (12.9)	-88.0* (18.6)	-47.7* (22.6)
Statistically significant at $p < 0.1$, but effect was in the unexpected direction			
Practice offered weekend visits	-41.8* (12.2)	-106.0* (34.0)	-64.3* (36.1)
Not statistically significant at $p < .01$			
Typical visit frequency for IAH beneficiaries	-37.5* (14.8)	-69.2* (26.3)	-31.7 (30.1)
Used formal risk stratification	-42.4* (15.2)	-60.1* (23.3)	-17.8 (27.8)
Percentage of primary care visits by telehealth or telephone in the first year of the COVID-19 pandemic	-53.7* (17.3)	-64.4* (24.0)	-10.7 (29.5)
Mode of hospital admission or ED visit notification	-65.2* (15.6)	-41.8* (17.2)	23.5 (23.2)
Performance on practices-reported quality measures used to calculate IAH payment incentive	-66.4* (27.2)	-41.7* (14.8)	24.7 (31.0)
Offered some beneficiaries at least one type of mobile diagnostic technology	-68.0* (19.3)	-37.5* (13.2)	30.5 (23.4)
Practice offered after-hours visits	-80.5* (27.7)	-42.8* (12.3)	37.6 (30.3)

Source: Medicare claims and enrollment data for 2010–2022 obtained from the Chronic Conditions Warehouse for IAH and matched comparison group beneficiaries in all IAH practices included in the evaluation.

Notes: See Exhibit E.1 for a description of the data sources used to measure care delivery activities.

* The difference is statistically significant at the 0.10 level.

ED = emergency department; SE = standard error.

Exhibit E.5. Number of practices that implemented most of the best practices, by structural characteristic

Structural characteristic	Number of practices	
	Implemented most of the best practices[a]	Did not implement most of the best practices
Organizational affiliation		
Independent	2	2
Academic health centers/health systems	4	3
HarmonyCares Medical Group[b]	0	5
Profit status		
For profit	2	7
Nonprofit	4	3
Participated in an accountable care organization		
Yes	3	6
No	3	4
Accepted Medicare Advantage plans		
Yes	6	6
No	0	2
Missing data[c]	0	2
Served at least 50% of its patients in an assisted living facility or other group residence		
Yes	1	2
No	5	8

[a] Sites that implemented at least four of the six best practices were considered as having implemented most of the best practices. The six best practices can be found in Exhibit E.4.

[b] HarmonyCares Medical Group was formerly known as Visiting Physicians Association before 2022.

[c] Data collection regarding whether practices accept Medicare Advantage occurred after two practices left the demonstration.

6. Limitations of the evaluation

6.1. Limitations related to analyses of the effects of IAH on spending and other outcomes

Evidence about causality from quasi-experimental evaluation designs is not as strong as evidence from a randomized controlled trial. Since the demonstration was not designed as a randomized controlled trial of the receipt of HBPC by beneficiaries or incentive payments by practices, the evaluation's performance-year and three-year analyses use a doubly robust difference-in-differences design with a matched comparison group. The causal validity of this design hinges on the parallel trends assumption, as we describe in Chapter 7 of Appendix C. This assumption is not directly testable because we cannot observe what would have happened absent IAH. Instead, we can assess several aspects of the design that would pose risks, such as large differences in characteristics between the IAH and comparison groups or large differences in outcome trends before the evaluation began; see Appendix C and Appendix D for detailed discussion of these issues. When presenting results, we systematically assessed the strength of evidence across several criteria, including whether differences in pre-period trends called a causal interpretation into question and whether qualitative evidence supported the regression result (see Chapter 7 of Appendix C for more details). However, it is not possible to identify and address all potential sources of confounding in a quasi-experimental design. Later in this chapter, we identify potential sources of confounding for the performance-year and three-year analyses.

Small numbers of participants could lead to imprecise results (large confidence intervals) and random fluctuations in estimated results. The small number of practices (at most 14) hampered the evaluation's ability to robustly measure the effects of the demonstration, especially if effects were small. There were a few reasons why the demonstration had a small number of participants. First, Congress limited the total number of beneficiaries in the demonstration, which restricted the number of practices that could participate. Second, practices withdrew from the demonstration over time, further decreasing the number of practices able to be evaluated. Third, the two- and three-year extensions of the demonstration did not allow enough time to recruit more practices to participate. With so few practices and beneficiaries, evaluation results could be subject to random fluctuations, which could lead to (1) larger deviations between the true and estimated effects of IAH and (2) wider confidence intervals, lessening the chance of an estimated effect being statistically significant.

Results are not generalizable to beneficiaries who do not meet IAH eligibility criteria, or to beneficiaries of other providers. IAH practices were not representative of other practices and clinicians that provide HBPC. Nearly all IAH practices offered primary care in the home only, not also in an office. Outside of IAH, many clinicians who deliver home visits do not deliver home visits exclusively. Nor were IAH beneficiaries similar to the average Medicare beneficiary. As we described in Chapter 1 of the report, IAH beneficiaries were much more disabled and much sicker. These factors suggest that results from this evaluation apply only to beneficiaries meeting IAH eligibility requirements who received HBPC from practices specializing in the delivery of HBPC.

Results during the COVID-19 pandemic are unlikely to inform what effects would be in non-pandemic years. The COVID-19 pandemic systemically disrupted beneficiaries' patterns of seeking care and clinicians' delivery of care. In sensitivity analyses, we found that COVID-19 diagnosis or hospital admissions generally did not affect the results, but COVID-19 illness may still have affected beneficiary disease severity, access to care, and feelings of isolation or loneliness. It also may have affected some providers through staffing shortages and disruption in care management activities.[47]

6.2. Limitations specific to specific analyses

The risk of confounding in the performance-year analysis from outside factors was higher in later years of the demonstration. The methods used for the performance-year analysis removed any *consistent* influence of unmeasured factors on outcomes (see Appendix C for details on the methodology). But it could not account for *changing* influence of unmeasured factors on outcomes. The further from a fixed baseline period that we measure changes in outcomes, the less confident we can be that no other factors may be affecting outcomes during that time. For example, over 10 years, IAH practices could have changed how they recruited new beneficiaries, which could have led to differences in the IAH population's health status over time that we could not measure in administrative data. These unmeasurable differences, and not the IAH payment incentive, could have driven differences in some outcomes. However, since estimated effects in the performance-year analysis were driven by results from Years 1 to 5 (July 2012 – 2017) which were closer in time to the baseline year (2011 – June 2012), this is probably not a major limitation.

The factors that lead a beneficiary to begin HBPC were likely complex, varied, and not measurable in the data sources available for the three-year analysis. These factors may have included a recent change in the beneficiary's health or functional status; physical or financial barriers the beneficiary faced when leaving home to receive health care; the beneficiary's preference for a more palliative or more aggressive style of health care; and whether the beneficiary felt comfortable having a health care provider make home visits. IAH practices cared for only a small share of IAH-eligible beneficiaries in the areas where those practices operated. If IAH practices cared for beneficiaries who differed in measurable ways from other IAH-eligible beneficiaries in the same market, there could have been differences in unmeasurable factors between the IAH and matched comparison groups. If those unmeasurable factors and their relationship to outcomes such as spending and hospital admissions were constant from the year before the index date through the end of the three years, then they do not pose a risk of confounding the results of the three-year analysis because of the difference-in-differences approach we used for the analysis. However, if those unmeasurable factors or their relationship to outcomes changed during the three years after starting HBPC, then they could have caused confounding. This limitation was most relevant to the three-year analysis because some beneficiaries in the intervention group in the baseline period of the performance-year analysis already had received HBPC from an IAH practice. Under the difference-in-differences design, as long as the proportion of beneficiaries who were not new to HBPC remained the same, those factors that may have led

[47] For more discussion of potential concerns related to the COVID-19 pandemic, refer to the evaluation report covering Year 7 of the IAH demonstration. For more information about results for the performance-year analysis during the COVID-19 pandemic, see IAH evaluation reports covering IAH Year 7, Year 8, and Year 9.

beneficiaries to begin HBPC would not be part of the estimated effects because they are differenced off as part of both the baseline and intervention periods.

The three-year analysis may have underestimated the effects of starting HBPC from an IAH practice in the second and third years. In the three-year analysis, we retained beneficiaries after their first year in the sample even if they were no longer receiving HBPC from an IAH practice. As a result, the effects in years two or three may have been smaller than they would have been if beneficiaries kept receiving HBPC from an IAH practice. As a sensitivity analysis, we estimated the effects of IAH on Medicare spending using an intervention group limited to beneficiaries who continued to receive HBPC from an IAH practice and their matched comparison beneficiaries. We found no statistically significant differences from the main results.

Our assessment of best practices during the demonstration is limited to information reported by IAH participants. Since we did not have data from other practices not participating in the demonstration, the results are limited to best practices among IAH practices. We also may not have captured all care delivery activities for each practice. To assess best practices, we used data from a worksheet that IAH practices completed and submitted to Mathematica, interviews with IAH practices, and practice-reported data to CMS, grouping practices according to use of specific strategies or characteristics and assuming that they were used throughout practices' participation in the demonstration. We did not collect complete data on two items included in the analysis (remote monitoring and mobile diagnostic technology), which could mean that we modestly undercounted practices performing these activities.

References

Brown, R. S., B. L. Carlson, S. B. Dale, L. Foster, B. R. Phillips, and J. L. Schore. "Cash and Counseling: Improving the Lives of Medicaid Beneficiaries Who Need Personal Care or Home and Community-Based Services." Final report submitted to the University of Maryland Center on Aging. Princeton, NJ: Mathematica Policy Research, August 2007.

Dale, S., S. Felt-Lisk, R. Shapiro, A. Ghosh, L. Moreno, M. Flick, and E. Lundquist. "Final Congressional Report on the Medicare Care Management Performance (MCMP) Demonstration: Incentivized Primary Care Practices Show Modest Changes in Care Management and Health IT." Princeton, NJ: Mathematica Policy Research, April 15, 2012.

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.

Massachusetts Department of Public Health and the Center for Survey Research, University of Massachusetts Boston. "Studying the Unmet Needs of Adults with Disabilities in Massachusetts, 2007." Boston: Massachusetts Department of Public Health, July 2008.

Mathematica Inc.

Our employee-owners work nationwide and around the world.

Find us at **mathematica.org** and **edi-global.com**.

Mathematica®
Progress Together