

**Systematic Review of CMMI
Primary Care Initiatives
FINAL REPORT**

**Prepared for:
Centers for Medicare and Medicaid Services (CMS)
Contract #: HHSM-500-2011-00008I
Deliverable #28
Project Officer: Patricia Markovich, Ph.D.**

Prepared by:
Kennell and Associates, Inc.
and its subcontractor:
RTI International

Report Release Date: February 2018



This page has been left blank for double-sided copying.

Systematic Review of CMMI Primary Care Initiatives Final Report

by

Kennell and Associates, Inc

Martin Cohen
Allison Russo
David Kennell
Sarah Irie

RTI International

Jim Derzon
Kevin Smith
Anupa Bir
Eva Chang
Kyle Emery
Will Parish
Amy Kandilov
Heather Kane
Shivani Reddy
Julia Cohen
Nikki Freeman
Jerry Cromwell*

Project Officer: Patricia Markovich, Ph.D.

Kennell and Associates, Inc
CMS Contract No. HHSM-500-2011-00008I

*Formerly with RTI International

This page has been left blank for double-sided copying.

Table of Contents

Executive Summary	1
1. Introduction.....	13
1.1 Design of the Systematic Review Study	14
1.2 Focus and Overview of This Report	15
2. Comparison of Initiative Features.....	17
2.1 Overview of the Six Initiatives	17
2.1.1 CMS-Convened Initiatives.....	17
2.1.2 State-Convened Initiatives	19
2.1.3 Non-Government Convened Initiatives	20
2.2 Commonalities and Differences among the Initiatives.....	21
2.2.1 Payer Participation	21
2.2.2 Patient Populations Covered.....	21
2.2.3 Implementation of Interventions.....	24
2.2.4 Practice Selection into the Initiatives.....	25
2.2.5 Payment Model	27
2.2.6 Beneficiary Attribution.....	28
2.2.7 Technical Assistance.....	28
2.2.8 Practice-Level Data Support.....	29
2.3 Conclusion	29
3. What are the Challenges and Opportunities in Convening Primary Care Initiatives?.....	31
3.1 Convener	31
3.1.1 Different Origins of Initiatives.....	31
3.1.2 Initiative Goals, Roles, and Responsibilities	33
3.2 Multi-Payer Structure.....	34
3.3 External Contextual Factors.....	36
3.3.1 Local and State Policies and Politics	36
3.3.2 Overlapping Health Care Reforms.....	37
3.3.3 Market Characteristics	38
4. What Facilitates Primary Care Transformation at the Practice Level?.....	39
4.1 Initiative-Level Facilitators.....	39
4.1.1 Financial Support	39
4.1.2 Learning System Features.....	43
4.1.3 Provision of Data to Practices.....	46
4.1.4 Easing Administrative Burdens Could Help Practice Transformation.....	49

4.2	Practice-Level Facilitators	50
4.2.1	Readiness	52
4.2.2	Health IT	54
4.2.3	Integrating New Staff.....	57
4.2.4	Identifying a Practice’s Patient Population.....	58
5.	How do the Initiative Impacts Compare and Contrast?.....	60
5.1	Differences in Outcomes by Initiative and Over Time.....	60
5.2	Differences in Outcomes by Setting.....	62
5.3	Summary and Discussion.....	64
6.	Using Pooled Regression Analysis to Examine Subgroups.....	65
6.1	Methodology.....	65
6.1.1	Primary Care Initiatives	65
6.1.2	Baseline Model: The Overall Impact of the Primary Care Initiatives	66
6.1.3	Triple Difference-in-Differences Models	66
6.2	Interpretation of the Baseline Model Results.....	67
6.3	Beneficiary Characteristics	68
6.4	Practice-Level Characteristics	70
6.5	County-Level Characteristics.....	75
6.6	Summary	75
7.	Meta-regression and QCA Analyses of Features that Contribute to High or Low Performance on Utilization and Expenditures	77
7.1	Meta-regression: Identifying Constellations of Features Associated with Performance on the Four Core Outcomes.....	77
7.1.1	Meta-regression Discussion and Conclusions	81
7.2	Qualitative Comparative Analysis (QCA).....	81
7.2.1	Analysis of Higher Performance.....	83
7.2.2	Analysis of Increases in Expenditures (Lower Performance).....	84
7.2.3	Key Lessons from QCA for High Performance.....	84
7.2.4	Limitations to Meta-Analysis and QCA	84
7.3	Summary and Discussion.....	85
8.	Evaluation Challenges	87
8.1	Length of Time for Transformation and Evaluation of Impacts.....	87
8.2	Defining Comparison Groups.....	87
8.3	Data Challenges	88

9.	Conclusions and Future Considerations.....	89
9.1	Summary of Key Findings	89
9.2	Conclusions.....	90
9.3	Implications for Scaling and Future Testing.....	92
9.3.1	Initiative Focus.....	92
9.3.2	Initiative Supports.....	93
9.3.3	Measures of Success	94

Table of Exhibits

ES-1	Overview of Characteristics of CMMI Primary Care Initiatives.....	3
ES-2	Summary of Key Findings.....	11
1-1	CMMI Study Questions.....	15
2-1	Overview of Characteristics of CMMI Primary Care Initiatives.....	22
2-2	Initiative-Level Support to Practice.....	23
3-1	Benefits and Challenges by Type of Convener.....	32
4-1	Annual Medicare PBPB Payments to Practices in Three Initiatives.....	41
4-2	Percentage of Settings that Required or Favored Practices with NCQA/State Accreditation.....	53
4-3	Extent to Which Settings Achieved health IT Transformation.....	56
5-1	Meta-regression Results Comparing the Four Core Outcomes with CPC, by Initiative (n = 22).....	61
5-2	Correlations of Setting Performance for Year 1, Year 2, and Year 3 Outcomes.....	62
5-3	Cumulative Three-Year Performance among 22 Primary Care Initiative Settings.....	63
6-1	Data Availability for the Pooled Regressions.....	66
6-2	Summary Statistics for Beneficiary Characteristics and Outcomes at Baseline.....	67
6-3	The Collective Impact of the CPC, FQHC, HCIA, and MAPCP Initiatives on the Four Core Claims-Based Outcomes Using Pooled Data.....	68
6-4	The Impact of the Initiative on Subgroups of Beneficiaries.....	69
6-5	The Impact of the Initiatives on Subgroups Based on Practice Characteristics.....	71
6-6	The Impact of the Initiatives on Subgroups Based on Practice-Level Patient Experience Scores.....	72
7-1	Meta-regression Predicting Medical Expenditures.....	78
7-2	Meta-regression Predicting Outpatient ED Visits— Model Including Percent Males.....	79
7-3	Meta-regression Predicting Outpatient ED Visits— Model Excluding Percent Males.....	80
7-4	Meta-regression Predicting Hospital Admissions.....	80
7-5	Meta-regression Predicting 30-day Readmissions.....	81
7-6	Characteristics of the Settings and Parameter of Fit for Successful Reduction in Cost.....	82
7-7	QCA Solutions and Settings with This Combination of Features.....	83

Executive Summary

Introduction to This Study

The Center for Medicare & Medicaid Innovation (CMMI) is statutorily authorized to test innovative health care payment and service delivery models that have the potential to lower Medicare, Medicaid, or CHIP expenditures while maintaining or improving the quality of beneficiaries' care. In the last decade, there has been a great deal of interest in improving primary care delivery and in the patient-centered medical home (PCMH) concept. Because there are still many questions about what it takes to successfully improve primary care and the impact of doing so, CMMI has launched a number of innovative models that aim to improve care, including six initiatives that were designed to test various advanced primary care models including PCMH. The PCMH concept was described in the Joint Principles of the Patient-Centered Medical Home (AAFP, 2007) and included six principles:

- A personal physician for each patient
- A physician-directed medical practice
- Coordinated care
- Enhanced access
- An emphasis on quality and safety
- Payments that would recognize the added value of PCMHs for patients

CMMI study questions. Over the course of this three-year study, we addressed eight CMMI study questions that can be grouped into four broad categories:

- What has been the impact of these primary care redesign initiatives on four core outcomes identified by CMS: Medicare costs, hospital admissions, emergency department (ED) visits, and 30-day readmissions? Have these impacts varied for different population groups?
- What external and internal factors assist practices with transforming into advanced primary care practices? What has been most important in helping practices meet transformation challenges?
- What have been the greatest challenges in evaluating the six initiatives' impacts?
- What are the implications of these findings for further testing or for scaling initiatives like these?

We conducted a meta-analysis of six CMMI initiatives to address these study questions. The six CMMI initiatives included in the study are the Comprehensive Primary Care (CPC) initiative, the Federally Qualified Health Center (FQHC) Advanced Primary Care Practice demonstration, the Independence at Home (IAH) demonstration, the Multi-Payer Advanced Primary Care Practice (MAPCP) Demonstration, the State Innovation Models (SIM) initiative, and the Health Care Innovation Awards Primary Care Redesign Programs (HCIA-PCR), which CMS identified as the most focused on primary care redesign.

We reviewed the formal evaluations of each of these initiatives, analyzed their findings, gathered data from the evaluations and the evaluators, and assessed the body of evidence they created as a whole and its implications for policy. Our goal was to identify the characteristics, supports, and internal and external factors related to transforming practices and to understand if particular aspects of such initiatives lead to better outcomes and lower cost. By comparing and synthesizing the experiences of the six heterogeneous initiatives and focusing on common outcomes across initiatives, we provide new findings and insights regarding factors that are associated with successful outcomes. We used two major analytic approaches to address these questions:

- **Qualitative analytic techniques** including content analysis of evaluation reports, interviews with key informants, and analysis of data we collected from the evaluators. We also used Qualitative Comparative Analysis (QCA), which allowed us to examine relationships between combinations of initiative characteristics and outcomes.
- **Quantitative analyses** including meta-analysis of effects from each evaluation and pooled regression analysis that used a patient-level dataset across multiple initiatives.

Commonalities and Differences between the Initiatives

Although the initiatives differed quite substantially on many dimensions (see *Exhibit ES-1*), they all sought to increase quality while reducing costs. The CPC and MAPCP initiatives were most similar: each used a multi-payer model, a per member per month (PMPM) payment model (though payment levels were more generous in CPC, and many MAPCP states also authorized payments directly to community care teams for care coordination), and offered similar kinds of technical assistance (TA) and data feedback to practices. A key difference is that CMS convened CPC, while the states convened MAPCP.

FQHC was similar to MAPCP and CPC with its clear focus on office-based primary care practice transformation to PCMH; however, this CMS-convened initiative was single-payer (Medicare) and provided lower payments to practices for transformation. Furthermore, the beneficiaries using FQHCs were more likely to be low-income.

IAH was unlike any other initiative in terms of its goals and focus on using a home-based primary care model to care for highly disabled, chronically ill beneficiaries. IAH was also the only initiative that did not include any requirements or assistance with respect to formal PCMH recognition. In addition, because the IAH practices were experienced in delivering home-based care, few changes in delivery were expected, and the IAH practices received no up-front financial support and little TA.

The HCIA-PCR initiatives also differed from the other initiatives in that they were not convened by CMS or the states. They focused exclusively on delivery innovation, as opposed to payment innovations, and no two cases were alike, though many shared a common focus on offering care management and coordination services.

The SIM states were largely building on the existing primary care infrastructure that already existed in those states. Most SIM states involved several components beyond primary care redesign and practice transformation, such as investment in health IT.

Exhibit ES-1. Overview of Characteristics of CMMI Primary Care Initiatives

	CPC	FQHC	IAH	MAPCP	SIM-Model Test Round 1	HCIA-PCR
Convener	CMS	CMS	CMS	State entity	State entity	Individual awardees
Participants	About 500 primary care practices operating in 7 geographical regions	About 500 entities serving a designated medically underserved area or medically underserved population providing primary care and serve 200+ Medicare beneficiaries	15 primary care practices, which provide home visits across 15 states to Medicare patients with chronic conditions/ disabilities	About 850 practices in 8 states (VT, ME, NC, RI, NY, PA, MI, MN); practices include some FQHCs, rural health clinics, and CAHs	6 states (AR, MA, ME, MN, OR, VT); some overlap with MAPCP (ME, MN, VT) and CPC (AR, OR)	14 awardee organizations; only 8 included in this study* (<i>mix of community-based organizations, providers, payers</i>)
Participating Payers	Multi-payer model—commercial insurers, Medicare, Medicaid, CHIP (4–8 payers per region)	Single-payer model—Medicare FFS	Single-payer model—Medicare FFS	Multi-payer model, which varied by state—Medicare, Medicaid, and commercial insurers (4–9 payers per state)	Varies by state—Medicaid in all states; multi-payer in 3 states; Medicare is not participating in any state	Varies by award, but not Medicare or Medicaid
Period of Performance	October 2012–December 2016	November 2011–October 2014	June 2012–September 2017	July 2011–December 2016	October 2013–January 2017	June 2012–June 2015 (some June 2016)
Distinguishing Features	<ul style="list-style-type: none"> Multi-payer model in 7 regions Generous financial support (\$20 initial PBPM from Medicare), payments from other payers (which gave practices a median payment of \$228,000 in 2013), intensive TA, and data provision Milestone approach to help practices transform 	<ul style="list-style-type: none"> Single-payer Low financial support (\$6 PBPM from Medicare, which provided \$31,400 per year per FQHC), although complemented by funding from HRSA and other sources Goal was to have FQHCs get NCQA PCMH recognition 	<ul style="list-style-type: none"> Home-based care model, not PCMH 15 experienced practices, so no up-front support and little TA from CMS Patients had multiple chronic conditions and ADL limitations 	<ul style="list-style-type: none"> Multi-payer model in 8 states Variation by states in model, financial support (average \$5.47 PBPM in direct payments from Medicare—averaging \$57,800 per practice per year—supplemented by indirect support and payment from other payers), and TA CMS joined ongoing initiatives in many states Practices had to have PCMH recognition on entry or within 6–18 months 	<ul style="list-style-type: none"> Goal is to transform the state’s health care system and have 80% of payments in each state under value-based or alternative payment models Multiple interventions in each state More emphasis on payment models and infrastructure (e.g., health IT/HIE) Medicare FFS not a participating payer 	<ul style="list-style-type: none"> Each award was a separate delivery reform intervention, some were multiple interventions Smaller initiatives, which were tailored to the specific local situations Lump-sum awards Little CMS involvement post-award

Abbreviations: FFS, Fee-for-Service; CHIP, Children’s Health Insurance Program; CAH, Critical Access Hospital; ADL, activity of daily living *The HCIA awardees included in this study are (1) PeaceHealth, (2) Denver Health, (3) Finger Lakes Health Systems Agency, (4) Pacific Business Group on Health, (5) TransforMED, (6) CareFirst, (7) Cheyenne-WyPCMH, and (8) Sanford Health

Challenges and Opportunities in Convening Primary Care Initiatives

Convener. The convener is the entity responsible for leading an initiative or setting and organizing its activities. Conveners include state agencies, CMS, commercial payers, and health care systems. The conveners decided how each initiative was designed, how primary care transformation was to take place, and which practices were included in each initiative.

Transformation initiatives that were largely run by the states rather than CMS were able to take advantage of local relationships and allow tailoring for local circumstances, but that implementation progress could be slow because states tend to place a high priority on maintaining or building within-state relationships and generally have limited resources in terms of staff and expertise. In addition, successful collaboration among multiple stakeholders requires conveners to communicate clearly the roles and responsibilities of the various payers and providers, as well as CMS's goals and strategies.

Multi-payer structure. A key design decision faced by conveners of an initiative is whether multiple payers participate. Single-payer initiatives may be easier to design and implement because design decisions do not require approval by other payers. However, a single payer only makes financial contributions and sets policy for only its share of the patients. Thus, it may be harder for a single payer to support transformation and to influence how practices transform care for all their patients.

Multi-payer initiatives are able to provide more financial resources to practices. However, each payer in a multi-payer structure may have different reporting requirements or quality measures. Key informants felt that multi-payer structures helped practices transform by increasing the financial resources to practices and the number of patients in a practice that were covered by modified care protocols. The benefits of a multi-payer structure are greatly increased if the payers can agree up-front on the reporting of a set of harmonized quality measures.

External contextual factors. External factors can have a large impact on implementation. For example, statewide initiatives can be affected by changes in state political offices, such as shifts in the governor's office. Limited staff resources at the state level also mean that implementation of statewide primary care initiatives may compete with implementation of other state priorities, such as implementation of state-run health insurance exchanges for the ACA.

What Facilitates Primary Care Transformation at the Practice Level?

The six initiatives allowed us to examine the challenges faced by practices in changing how they delivered primary care and how the challenges were overcome. We categorized these issues into (1) the ways that initiatives tried to support practices in confronting the challenges of transforming to advanced primary care practices and (2) the features of the practices themselves that were challenges and facilitators for transformation.

Initiative-Level Supports to Practices

Financial support. Practices need financial support for transformation because Medicare's customary fee-for-service payments may not provide reimbursement for some

important advanced primary care activities (e.g., coordination and reporting). Conveners must select the level and form of financial support that best supports practice transformation given (1) the current level of transformation of the practice, (2) the requirements of the initiative, and (3) other resources that a practice can use to pay for transformation activities.

The initiatives provided financial support either directly to practices through care management fees (typically PBPM payments) or indirectly through such mechanisms as shared savings or pay-for-performance (P4P) incentives. The levels of support provided by Medicare to each practice varied from about \$31,000 per year in FQHC (which most FQHCs thought was insufficient for PCMH implementation) to \$146,000 per year in CPC (which most CPC practices thought was adequate for implementing CPC). That the FQHC initiative was a single-payer initiative and CPC was a multi-payer initiative may also have influenced practices' attitudes about the adequacy of financial support. The median payment per practice in CPC was \$228,000 in 2013 when considering all payers. Despite the perceived lack of sufficient funds, about 70 percent of initiative FQHCs achieved NCQA Level-3 PCMH recognition by the end of the initiative, with nearly half reaching that goal in the final three months of the demonstration.

Shared savings can motivate practices to be attentive to reducing costs and utilization by providing payments to practices that meet a specified savings target while meeting specified quality standards. To overcome some of the challenges of shared savings, some initiatives and settings (CPC and MAPCP-PA) provided both shared savings and PBPM payments. In the early years, the non-refundable PBPM payments supported capital and labor investments; these PBPM payments were reduced in later years when CMS added shared savings to incentivize additional improvements in efficiency and quality.

Learning system features. To achieve advanced primary care, many practices were challenged in adapting their electronic health record (EHR), hiring new staff, and integrating new staff and functions into their workflow. The six initiatives used a variety of learning system approaches to providing TA and other educational support to initiative settings. Some adopted a centrally managed learning system enabling consistent national messaging across large, multi-site initiatives. Others adopted distributed systems, which allowed tailoring of TA to the circumstances of different regions or individual practices.

Delivery modes for TA also varied with initiatives typically providing multiple forms of TA, including online or in print, conference calls, webinars, peer-to-peer learning collaboratives, and one-on-one coaching of individual sites. Practices generally preferred the more tailored forms of TA (one-on-one coaching and peer-to-peer learning). Our attempts to test the relationship between TA and outcomes were inconclusive. In our meta-regression analysis, some form of TA significantly contributed to impacts on three of the four outcomes (Medicare expenditures before initiative fees, hospital admissions, ED visits, and readmissions), although the form varied by outcome. In contrast, our QCA analysis found that settings that performed better on Medicare cost savings actually had lower practice participation in TA activities (as judged by evaluators for each setting overall) than practices in lower-performing settings. This QCA finding may underscore the need for more tailoring of TA based on differences in practices' level of sophistication. More advanced practices sometimes viewed the available TA activities or resources as too basic to be useful to them, while the most tailored (and resource-

intensive) form of TA—one-on-one direct assistance—often was offered predominantly to lower-performing practices that needed the most help.

Practices received TA on a wide variety of topics across the initiatives, including PCMH recognition requirements, leadership and organizational change, practice-level quality improvement, primary care staff development and integration of new staff, care coordination, health IT, and use of practice feedback data. Key informants repeatedly mentioned the importance of starting TA early, tailoring TA topics to recipients' levels of expertise, and that TA topics can change over time as practices make progress on transformation. Both key informants and evaluation reports identified health IT as a central TA topic because many practices faced challenges in using their EHRs to obtain near real-time knowledge of discharged inpatients and emergency department visits. Key informants also noted that a well-organized, user-friendly website helped practices navigate the learning system's offerings and materials.

Provision of data to practices. Practices need patient-level data to coordinate care for high-risk patients and practice-level data on quality, utilization, and costs to deliver care more efficiently and to monitor their performance. Conveners anticipated these needs and typically provided both practice-level and patient-level data to practices. Many practices found some value in these reports, but usage varied greatly across initiatives. The evaluation reports and our key informant interviews highlighted a range of challenges limiting the value of the data:

- Quarterly, claims-based data were too infrequent to provide actionable data for care management.
- Some felt feedback reports were not actionable because results could be driven by factors beyond their influence, such as the behavior of specialists and hospitals (Peikes, Anglin, et al., 2016).
- Privacy barriers to sharing mental health and substance abuse (MH/SA) data posed a content problem for feedback reports.
- Practices wanted aggregated, multi-payer data.
- Reviewing external data reports was not part of a physician's normal workflow.
- Practices did not receive enough training in how to interpret and use the data provided.

Easing administrative burdens could help practice transformation. Some participating practices felt burdened by requirements associated with the initiatives, including (1) the level of documentation and reporting required, (2) the effort involved in determining which patients were eligible for enhanced services under the initiative (e.g., care coordination or home services), and, (3) in the case of CPC, the number of milestone requirements they were expected to achieve.

Practice-Level Facilitators

We identified some important factors that were internal to practices and helped them become advanced primary care practices. These factors included practice readiness, having health IT capabilities, and the ability to integrate new staff into the practice's workflow.

Practice readiness. Practices' prior transformation experience facilitated initiative implementation and that effective and committed practice-level leadership helped build a collaborative work environment and support implementation. However, smaller practices needed additional support because they may have lacked the staff, resources, and capacity to manage implementing transformation.

Health information technology. We also found that health IT was an important factor in transformation and that difficulty in implementing an EHR was a frequent barrier to broader practice transformation activities. Health IT solutions needed to evolve to meet the demands of practice transformation, including improved care coordination, care management, and payment reform.

Integrating new staff. Most practices needed to add staff to become advanced primary care practices. Allowing sufficient time for staff to perform practice transformation activities and preparing for staff turnover each facilitated practice transformation. Our study found that having clearly defined roles for newly hired staff improved hiring practices and retention, and facilitated practice transformation.

Impact of the Initiatives

The evaluation reports, the analyses we conducted for this study using the Evaluator Summary Form (ESF) survey instrument developed for this project, and our interviews with key informants all indicate that initiative practices made large strides toward becoming PCMHs or advanced primary care practices. For example, while less than 10 percent of initiative FQHCs had any PCMH recognition status prior to the initiative, 70 percent of initiative FQHCs achieved NCQA Level-3 recognition by the end of the initiative. In an exploratory analysis we conducted for this meta-evaluation using data from the ESF, we found that across the initiatives, evaluators' judgments about the practices they studied at the setting level indicated that practices had increased their practice transformation score by an average of more than 60 percent during the initiatives.

We examined the impact of four of the six initiatives combined over a cumulative three-year period for the four core Medicare outcomes identified by CMS (Medicare expenditures before initiative fees, outpatient ED visits, hospital admissions, and 30-day readmissions). We found that when considered collectively, the results were not statistically significantly different from zero at a significance threshold of $p < 0.10$ for each of the four core outcomes relative to a comparison group. When we compared the results for the four core outcomes for 22 initiative setting levels (eight MAPCP states, seven CPC regions, six HCIA awards, and FQHC) using a significance level of $p < 0.10$, we found variation among the 22 settings (see *Exhibit 5-3*):

- Four of these 22 settings had lower Medicare expenditures, while four had higher expenditures.
- Four of these 22 settings had lower outpatient ED visits, while five settings had higher outpatient ED visits.
- Four settings had lower hospital admissions, and three had higher hospital admissions.

- Two settings had lower 30-day hospital readmissions, and two had higher rates.

Overall, across the 22 settings and four outcomes summarized (88 possible results), 68 percent of the results showed no effect, 16 percent showed improvements in outcomes, and 16 percent experienced worse outcomes, relative to comparison groups at the $p < 0.10$ level.

HCIA TransforMED, a health IT intervention, was the only setting to show improvement across all four core outcomes at the $p < 0.10$ level. Two CPC settings (OK and OR) each showed improvement for at least two of the four outcomes at $p < 0.10$.

Impact on Population and Practice Subgroups

Primary care initiatives may have more (or less) favorable impacts on particular subgroups of Medicare beneficiaries, on practices with certain characteristics, or in health care markets with certain characteristics. This systematic review gave us the opportunity to understand which populations and which practices experienced the largest impact from primary care initiatives. We were able to do this by combining beneficiary-level claims data across four initiatives (CPC, MAPCP, FQHC, and HCIA-PCR) to determine the differential effects of primary care initiatives on different groups in terms of the four core outcomes. We used regression analysis with the pooled data to estimate the impacts of the primary care initiatives on beneficiary and practice subgroups.

Impact on beneficiaries with greater health needs. Because many of the initiatives focused on care management and care coordination activities for beneficiaries with chronic conditions, we expected to see more favorable impacts of the primary care initiatives on beneficiaries with more health care problems and higher health care needs because these beneficiaries might benefit more from the greater care coordination and other primary care initiatives aimed at providing more efficient care. For example, this more favorable impact was found in the evaluation of the MAPCP Michigan and New York initiatives.

We tested whether four of the initiatives collectively had a significant impact on beneficiaries with greater health needs (i.e., those who had HCC scores in the top quartile in the baseline period, those who originally qualified for Medicare due to disability, and those who were dually eligible for Medicare and Medicaid) and whether the collective impact of the four initiatives was significantly larger or smaller for these beneficiaries.¹ Specifically, we found the following:

- Among beneficiaries in the highest quartile of HCC scores, the growth in total Medicare expenditures among intervention beneficiaries was \$49.39 PBPM less than beneficiaries in the highest quartile of HCC scores in the comparison group ($p < 0.10$).
- Among beneficiaries who originally qualified for Medicare due to disability, the overall growth in total Medicare expenditures among beneficiaries in the intervention group was \$38.60 PBPM less than that of similar beneficiaries in the comparison group ($p < 0.10$).

¹ FQHC, CPC, MAPCP, and six HCIA awardees.

- Among beneficiaries who did *not* qualify for Medicare due to disability, the initiatives collectively decreased the overall probability of having any readmission within 30 days of a hospital discharge by 0.43 percentage points, relative to non-disabled Medicare beneficiaries in the comparison group ($p < 0.10$).
- Among non-dually eligible beneficiaries, the initiatives collectively decreased the overall probability of having any readmission within 30 days of a hospital discharge by 0.42 percentage points relative to non-dually eligible beneficiaries in the comparison group ($p < 0.10$).

The slower growth in Medicare expenditures for the beneficiaries in the highest quartile of HCC scores and for beneficiaries who originally qualified for Medicare through disability is consistent with our prior expectation about a stronger impact of these initiatives for beneficiaries with high health care needs, although we did not find a similar effect for the other three core outcomes.

Impacts on practices. We also found impacts of different types of primary care practices on outcomes. We found the four initiatives collectively slowed growth in Medicare expenditures and lowered rates of inpatient admissions and ED visits among practices with fewer than six practitioners ($p < 0.10$) and among those that were primary care-only practices (i.e., not multispecialty practices), relative to practices with these same characteristics who did not participate in the initiatives.

We also identified practices that scored highly on five composite measures from the overlapping elements of the practice surveys, limited only to CPC and MAPCP practices and CPC comparison practices with responses to the practice survey. We found the following:

- Practices with high care management scores and high quality indicator scores saw slower growth in Medicare expenditures and lower rates of hospital admissions, relative to comparison practices with these characteristics ($p < 0.10$).
- Practices with high care coordination scores also saw lower rates of hospital admissions, relative to comparison practices with similar characteristics ($p < 0.10$).

Initiative Features Related to Higher Performance

To identify features of the initiatives related to higher performance, we used both meta-regression and Qualitative Comparative Analysis (QCA). The meta-regression allowed us to analyze factors associated with higher performance on each of the core outcomes separately. For the QCA, we defined higher performance based on improvement on the total cost of care. These approaches are complementary, and the results obtained from each provide insights into how settings achieved success. We reached three key conclusions:

- Provision and receipt of TA can promote better performance, but practices' interest in TA and the benefit they perceive depends in part on how well the TA is tailored for differences in practices' level of sophistication.
- Feedback reports can improve performance but only to the extent practices actually use the data, which varied across initiatives and settings.

- Adequate financial support allowed the hiring of new staff, and, in the QCA, financial and staff adequacy were present for high performers who did not use feedback reports. Having the ability to hire new frontline staff was also associated with reduced hospital admissions in the meta-regression.

Evaluation Challenges

In estimating the impacts of each initiative, evaluators had to confront many challenges, including the following:

- **Length of the evaluations.** Key informants thought that three to four years was a sufficient length of time for primary care practices to make changes to their care delivery methods, but that this length of time was sometimes insufficient to observe the impacts of these changes on the four core outcomes, particularly in the initiatives in which much of the transformation activity occurred near the end of the initiative.
- **Defining comparison groups.** Because of the lack of randomized designs, all the evaluations had to create comparison groups and all evaluators thought that identifying valid comparison groups was challenging.
- **Data challenges.** For initiatives that included Medicaid patients, access to Medicaid data was delayed and many evaluations could not use Medicaid data. Evaluators were also challenged by the level of practice consolidation during the initiative period and the confounding effect of other CMS initiatives on their impact analyses.

Implications for Scaling and Future Testing

Our review of the six initiatives points to a number of implications on scaling and further testing. We categorized these implications into four areas:

- **Initiative focus.** Future initiatives could be focused on practices at specified levels of readiness. Small targeted initiatives could also be used to provide smaller-scale test of concepts prior to full-scale testing, as in the HCIA awards.
- **Initiative supports.** We found that initiatives features like TA, data support, and financial support were important in helping practices transform but that it was not possible to isolate the optimal amount and form of each support. The design and evaluation of future initiatives could help CMS refine its understanding of the necessary levels of different supports for transformation.
- **Financial support.** In future testing there could also be value in varying specific types of financial support based on specific practice characteristics (e.g., perhaps varying the level of PBPM support among smaller and larger practices, providing more up-front support for practices that are not yet PCMHs, or testing the impact of shared savings versus PBPM payments for larger practices only).
- **Measures of success.** Having a longer time period for future initiatives would be useful for initiatives that focus on practices that are less ready to transform.

Exhibit ES-2 provides a summary of the key findings.

Exhibit ES-2. Summary of Key Findings

1. What has been the impact of these primary care redesign initiatives on the four core outcomes? Have these impacts varied for different population groups?

- Practices made substantial progress transforming to advanced primary care practices.
- The collective impact of four of the six initiatives over a cumulative three-year period (CPC, MAPCP, FQHC, and six HCIA awards) was not significantly different from zero at a threshold of $p < 0.10$ for the four core outcomes (Medicare expenditures before fees, outpatient ED visits, hospital admissions, and 30-day readmissions) relative to a comparison group.
- When the impacts were analyzed at the more granular level of the 22 initiative setting levels (eight MAPCP states, seven CPC regions, six HCIA awards, and FQHC), we found some positive and some negative results for the four core outcomes. Almost half of these 22 settings improved on at least one core outcome ($p < 0.10$), but the improvements were generally modest, and only three settings improved on at least two core outcomes.
- We pooled the data from four initiatives and found that, collectively, the initiatives had decreased the growth in Medicare expenditures before fees ($p < 0.10$) for beneficiaries in the highest quartile of HCC scores and for beneficiaries who originally qualified for Medicare due to disability. We also found decreased growth in Medicare expenditures before fees ($p < 0.10$) for practices with fewer than six practitioners and primary care-only practices.

2. What external and internal factors assist practices with transforming into advanced primary care practices? What has been most important in helping practices meet transformation challenges?

- Multi-payer structures aid practice transformation, especially if performance measures are harmonized across payers.
- Practices that are less advanced, especially smaller practices, may need higher up-front financial support.
- TA needs to start early, but mid-course adjustments are often needed; tailoring for practice level of transformation is desirable; health IT and use of data are important TA topics.
- While many practices perceived some benefit from the feedback data they received, there were many challenges in using these reports.
- Practices are at various stages of readiness to transform; some internal supports, including a practice champion and high-functioning health IT, can greatly aid transformation.

3. What were the greatest challenges in evaluating the initiatives?

- Insufficient time to observe impacts of the initiatives
- Identifying valid comparison groups
- Data challenges, especially delays in Medicaid data

4. What are the implications of these findings for further testing?

- More testing could be done on initiative focus (such as focus on certain features or levels of practice readiness), initiative supports (including financial supports), and having longer initiatives.

This page has been left blank for double-sided copying.

1. Introduction

Section 1115A of the Social Security Act authorizes CMMI to test innovative health care payment and service delivery models that have the potential to lower Medicare, Medicaid, and CHIP expenditures while maintaining or improving the quality of beneficiaries' care (42 U.S.C. 1315a). Under the law, preference is to be given to models that improve coordination, efficiency, and quality.

CMMI is interested in improving primary care delivery and in whether delivery models like the patient-centered medical home (PCMH) and other advanced primary care concepts can improve quality and reduce costs. This interest builds on long-standing questions about how to improve primary care delivery. In 2007, in an attempt to provide an improved primary care model, four groups of primary care physicians published the Joint Principles of the Patient-Centered Medical Home (AAFP, 2007). The six principles described the core features of PCMHs:

- A personal physician for each patient
- A physician-directed medical practice
- Coordinated care
- Enhanced access
- An emphasis on quality and safety
- Payments that would recognize the added value of PCMHs for patients

In 2008, the National Committee for Quality Assurance (NCQA) established requirements for PCMH recognition, which are widely used to assess primary care delivery (and were revised in 2011 and 2014). Others such as the Accreditation Association for Ambulatory Health Care (AAAHC) and The Joint Commission have since created their own accreditation programs.

Although many studies have assessed primary care redesign, there are still many unanswered questions about how best to promote primary care transformation and whether such efforts lead to improved outcomes (Hoff, Weller, & DePuccio, 2012; Peikes et al., 2012; Jackson et al., 2013; Edwards, Bitton, Hong, & Landon, 2014; Landon, 2014; Wiley et al., 2015; and Sinaiko et al., 2017). For example,

- What type of support is necessary to overcome the challenges of practice transformation? Is the key factor some type or level of financial support, is it a certain type and level of technical assistance/data support, or does it take a combination of these elements?
- What other factors are associated with successful practice transformation? Are there internal organizational factors that are necessary? Are there certain external policy/regulatory factors that facilitate successful transformation?

- Does practice transformation lead to better outcomes in terms of higher quality care, better patient experience, and lower costs? What are the intervention characteristics, implementation processes, and other factors associated with better outcomes?

To try to address such questions and to test whether advanced primary care models could increase quality and reduce costs, CMMI launched six initiatives in 2011-2013. These six initiatives are described in Chapter 2 and include the following:

- Federally Qualified Health Center (FQHC) Advanced Primary Care Practice demonstration
- Comprehensive Primary Care (CPC) initiative
- Independence at Home (IAH) demonstration
- Multi-Payer Advanced Primary Care Practice (MAPCP) Demonstration
- Primary care models included in the State Innovation Models (SIM) initiative
- Health Care Innovation Awards Primary Care Redesign Programs (HCIA-PCR)

1.1 Design of the Systematic Review Study

CMMI specified eight broad questions for a **systematic review** about the six initiatives that they hoped could be addressed in this study (see *Exhibit 1-1*). Our study builds upon the CMMI evaluations of each initiative to address these over-arching questions using a meta-evaluation approach. We reviewed the formal evaluations of each of the six initiatives, analyzed their findings, gathered data from the evaluations and the evaluators, and assessed the body of evidence they created as a whole and its implications for policy.² Our goal was to identify the initiative features and other factors related to transforming practices and to understand if particular aspects of practice transformation lead to better outcomes and lower cost. By comparing, contrasting, and synthesizing the experiences of the six heterogeneous interventions and focusing on common outcomes across initiatives, we hope to provide new findings and insights regarding factors that are associated with successful outcomes and primary care models that are particularly successful with specific subgroups like the chronically ill or dually eligible beneficiaries.

We used two major analytic approaches to address these questions:

- **Qualitative analytic techniques** such as content analysis of evaluation reports, analysis of interviews with key informants, and analysis of data we collected from the evaluators to provide a detailed understanding of the intervention components and implementation contexts as well as insight on crosscutting themes. We also used Qualitative Comparative Analysis (QCA), which allowed us to examine relationships between combinations of conditions in the initiatives (i.e., features, characteristics) and outcomes.

² Our study is a meta-evaluation rather than a formal systematic review because it is not an explicit search of the literature that includes or excludes certain studies based on specified criteria.

- **Quantitative analyses** including meta-regression and meta-analysis of effects from each evaluation. These analyses include statistical techniques for combining, comparing, and evaluating the treatment effects of similar interventions and allow generalizations to be made about the evaluations with similar outcome measures. We also conducted our own regression analysis using a pooled, patient-level dataset across multiple initiatives to estimate the collective impact of the initiatives for a variety of sub-populations and types of primary care practices.

Exhibit 1-1. CMMI Study Questions

1. What are the commonalities among and differences between the initiatives, including the following?
 - Structure (e.g., multi-payer vs. single-payer, state-led vs. CMS-led) and scope
 - Participation (e.g., payers, practices, states, patients)
 - Payment method
 - Site selection and generalizability of the selected sites to the expansion population
 - Beneficiary attribution
 - Implementation of interventions
 - Technical assistance
 - Provision and use of data
2. How do the evaluation findings compare and contrast?
3. Were there any common subgroups (e.g., beneficiaries with multiple chronic conditions, dual eligible) across evaluations that experienced differential impacts? How did their results differ?
4. What are some commonalities in features (e.g., model structure, convener, and payment support) and external contextual factors among
 - high-performing initiatives and settings within the initiatives?
 - low-performing initiatives and settings within initiatives?
5. What challenges did conveners (CMS, states, or awardees) encounter in designing and implementing the initiatives? What lessons were learned about addressing these challenges?
6. What challenges did practices face in these initiatives and how they were overcome? What initiative features best supported transformation including payment, learning system models, and data sharing?
7. What challenges did CMS and the evaluators encounter in evaluating the initiatives?
8. What are some additional important considerations for expanding initiatives like these at a larger scale? In future primary care initiatives, are there other features worth testing?

1.2 Focus and Overview of This Report

We organized this report according to the eight major research questions in *Exhibit 1-1*. Chapter 2 describes the commonalities and differences among the six initiatives. Chapter 3 focuses on initiative design issues, such as the type of convener and whether the initiative is supported by a single payer or multiple payers. Chapter 4 discusses practice transformation from two perspectives: (1) what are the facilitators that can be provided by initiatives to support practices, and (2) what characteristics of practices best support transformation to an advanced

primary care practice? Chapters 5, 6, and 7 analyze the impacts of the initiatives on outcomes, including their impacts on population and practice subgroups and the features that were common to high-performing practices. Chapter 8 then discusses the challenges in measuring the impacts of the initiatives. Chapter 9 presents conclusions and discusses some of the implications for future scaling of the initiatives. *Appendices A* and *B* describe the primary and secondary data sources and analytic approaches used in this meta-evaluation.

2. Comparison of Initiative Features

2.1 Overview of the Six Initiatives

The six initiatives used different approaches to attempt to increase the quality of care while reducing or maintaining its cost. After a brief overview of each initiative and a high-level comparison in *Exhibit 2-1*, the remainder of this chapter describes in more detail the differences and similarities among the six initiatives across multiple dimensions to address study question 1, which asks: What are the commonalities among and differences between the initiatives, including

- Structure and scope,
- Participation,
- Payment method,
- Site selection and generalizability of the selected sites,
- Beneficiary attribution,
- Implementation of interventions,
- Technical assistance, and
- Provision and use of data?

Chapters 3 and 4 then discuss lessons learned about key initiative features as well as practice-level factors that are important to primary care transformation.

One difference among the six initiatives is the *convener* or leader of the initiative. The convener is the entity responsible for organizing activities at the initiative level. CMS was the convener for the FQHC, CPC, and IAH initiatives, while the respective states convened the MAPCP and SIM initiatives. The individual organizations that received the HCIA awards were the HCIA conveners. Conveners make many important decisions about the structure and implementation of initiatives. Different types of conveners have various forms of authority to use policy levers or purchasing power to influence provider and payer participation in an initiative. Chapter 3 discusses lessons learned about different types of conveners.

2.1.1 CMS-Convened Initiatives

Comprehensive Primary Care (CPC) initiative. CPC was a four-year multi-payer collaboration convened by CMS with more than 35 public and private health care payers and roughly 500 participating practices in seven geographic regions. CPC's goal was to improve primary care delivery and achieve better care, smarter spending, and healthier people while enhancing patient and provider experience. In CPC, participating practices were required to meet a series of annual milestones on their paths to implementing five key CPC functions: (1) risk-stratified care management, (2) access and continuity, (3) planned chronic and preventive care, (4) patient and caregiver engagement, and (5) coordination across the medical neighborhood. Although these five functions are similar to PCMH functions, achieving NCQA Level-3 recognition was not a goal of CPC.

CPC is distinguished from the other initiatives by the support it provided to practices. To help participating practices implement the CPC milestones, the CPC initiative provided three types of support: (1) financial support in the form of an enhanced, non-visit-based payment from participating payers as well as the opportunity to receive shared savings; (2) learning activities and technical assistance sponsored by CMS; and (3) data feedback on each practice's progress in improving patient outcomes and controlling costs, provided quarterly by CMS for Medicare fee-for-service (FFS) patients and with varying periodicity by many of the other participating payers. Further, because CMS favored selecting practices that used electronic health records (EHRs) and that were accredited as PCMHs, initiative practices were relatively advanced.

Federally Qualified Health Center (FQHC) Advanced Primary Care Practice demonstration. CMS convened this three-year demonstration to support the transformation of approximately 500 FQHCs in delivering advanced primary care to Medicare beneficiaries. As part of transforming their delivery of primary care, FQHCs were expected to obtain Level-3 PCMH recognition (2011 standards) from the National Committee for Quality Assurance (NCQA) by the end of the demonstration. FQHCs in the initiative received financial support, technical assistance, data, and feedback reports to support the transformation of FQHCs into PCMHs.

Three factors distinguish this initiative. First, it focused on FQHCs, which serve low-income, vulnerable, populations in underserved areas. Second, FQHCs were at an early stage of practice transformation entering the demonstration, as indicated by the fact that only 7 percent of FQHCs had any PCMH recognition at the start of the initiative (in contrast to about 40 percent of the CPC practices). Third, the FQHC initiative was a single-payer initiative, and Medicare provided a relatively low level of financial support for a small subset of FQHC patients.

Independence at Home (IAH) demonstration. The Affordable Care Act required that CMS conduct a demonstration to test the ability of a home-based primary care delivery and incentive model—led by physicians or nurse practitioners (NPs)—to reduce expenditures and improve health outcomes of high-need Medicare beneficiaries. The 15 practices participating in this ongoing five-year demonstration were expected to design and implement coordinated care plans that were tailored to individual beneficiaries' chronic conditions and responsive to their preferences. If those efforts resulted in lower Medicare expenditures, the practices that also met specified quality measures received a portion of the Medicare program's savings. IAH differs markedly from the other initiatives in five ways:

- The initiative was targeted at very disabled beneficiaries (among other requirements, individuals had to have at least two chronic conditions and need human assistance with two or more activities of daily living).
- The IAH model focuses on home-based care provided by physicians/NPs.
- Because the practices were already using the home-based primary care model at the start of the demonstration and did not anticipate making changes to their care delivery models, they received no up-front financial assistance to transform, and the technical assistance from CMS was limited primarily to helping with the initiative's reporting requirements.

- The only financial support was through potential shared savings.
- Authorizing legislation, rather than by the convener, detailed initiative requirements.

2.1.2 State-Convended Initiatives

Multi-Payer Advanced Primary Care Practice (MAPCP) Demonstration. In this demonstration CMS joined state-sponsored PCMH initiatives already underway in eight states. The PCMH initiative in each state was convended by a state agency as part of a state-sponsored reform initiative. Medicaid and private health plans participated in addition to fee-for-service Medicare. Practices and other entities (e.g., community health teams) were provided with financial support, technical assistance, and feedback reports to support practice transformation and provision of PCMH services.

MAPCP differed from some of the other primary care redesign initiatives in three other ways:

- **Diversity of practice transformation goals.** Although the state initiatives were required to promote the principles of “advanced primary care practice,” each state convener could adopt its own definition of what constituted such practice. Although all but two of the MAPCP states (Michigan and Minnesota) elected to define “advanced primary care” in alignment with the NCQA’s recognition standards, many of the MAPCP states added additional expectations for practices to reflect local priorities.
- **Integration of community-based resources to support advanced primary care practices.** Six states (Maine, New York, North Carolina, Michigan, Rhode Island, and Vermont) funded community health teams (CHTs), community-based practice support networks, or physician organizations to involve community-based resources.
- **PCMH recognition.** Five of the eight MAPCP states required PCMH recognition at the start of the initiative and the other three states required it within 6–18 months.

The MAPCP initiative is most similar to CPC in that it provided financial support through multiple payers and technical assistance to practices, most of which had already started to transform. However, it differed from CPC in that each state had slightly different primary care redesign goals, each state provided financial support and technical assistance somewhat differently, each state implemented its primary care redesign changes without the milestone approach used in CPC, and each state was required to have consistent payment methods applied by all payers in a state.

State Innovation Models (SIM) initiative. In the six Round 1 model-test states, the SIM initiative tested the ability of state governments to accelerate statewide health care transformation. Two common aims among the six SIM Round 1 Test states were to shift the state’s health system from encounter-based service delivery to care coordination, and from volume-based to value-based payment mechanisms. The underlying belief was that better coordinated and more accountable health care would lead to higher quality care at lower total cost and, ultimately, to improved population health. The SIM initiative recognized the unique role states can play—as purchasers and regulators—to bring about or hasten health care

transformation using policy and regulatory levers to engage a broad range of stakeholders and to build on existing efforts.

The states took different approaches with their SIM awards to achieve this transformation. To varying degrees, states focused on models that emphasized two approaches: (1) primary care practice transformation to patient-centered, coordinated care and (2) integration of primary care providers with providers of acute care, behavioral health services, and long-term services and support. All six states tested one or more of four major delivery system and payments models:

- PCMHs (four states)
- health homes for medically complex populations established under Section 2703 of the Affordable Care Act (four states)
- integrated or accountable care systems (five states)
- episodes-of-care payment models (two states)

To facilitate transformation, the states are using a number of enabling strategies, including the following:

- practice transformation facilitation
- workforce development
- health IT investment
- data analytic capacity building
- stakeholder engagement (including consumer education)

SIM is testing whether transformation of the state's health care system can be successful when implemented within the context of a broad *state* plan owned by key public and private stakeholders. This initiative varies substantially from the MAPCP, CPC, and FQHC initiatives because SIM is attempting to transform each state's health care system, rather than just having practices become PCMHs or advanced primary care practices. In addition, Medicare was not a participating payer in SIM, and CMS did not provide technical assistance to practices (CMS provided TA to the states, and states provided TA to practices).

2.1.3 Non-Government Convened Initiatives

The Health Care Innovation Awards Primary Care Redesign Programs (HCIA-PCR). CMS made awards to 14 primary care redesign programs, which varied widely in terms of their intervention and market characteristics. There were a variety of conveners in HCIA: some programs were led by health care systems, some by commercial payers, and some by public/private partnerships. We reviewed a subset of eight HCIA-PCR awardees that CMS thought were most focused on primary care redesign. Although these eight awardees used different approaches to improve primary care delivery, all eight implemented care management, seven implemented care coordination, and six awardees focused on providing patient-centered care. The eight HCIA-PCR awardees had a narrower geographic scope than many of the other

initiatives. Two of the eight awardees attempted to transform the delivery of primary care in local health care markets only, and one focused on regional markets within a single state. Two other awardees attempted to change the way in which primary care services were delivered statewide, and three tried to redesign primary care delivery systems in more than one state. Many of the awards had a narrow focus for their intervention. For example, HCIA TransforMED aimed to assist primary care practices upgrade their health IT capabilities so that they could manage the health of their patient populations. Other awards focused on practice transformation and care coordination in remote island communities in Alaska (HCIA-PeaceHealth) and practice transformation at a safety net health system (HCIA-Denver Health). The HCIA awards differed from the other initiatives in that they were generally smaller projects and tested different features of practice transformation (e.g., high-risk identification in HCIA CareFirst or health IT in HCIA TransforMED) or for specific populations (rural care in HCIA-PeaceHealth and HCIA-Cheyenne).

2.2 Commonalities and Differences among the Initiatives

To address study question 1, this section describes the commonalities and differences in important features of the six initiatives. *Exhibits 2-1* and *2-2* present a summary comparison, while we present more detail below.

2.2.1 Payer Participation

The initiatives varied by (1) whether they were single-payer or multi-payer initiatives and (2) whether Medicare or Medicaid was a payer. A single-payer initiative is one in which only the convener provides financial support. FQHC and IAH are examples of this. Multi-payer initiatives, such as CPC and MAPCP, brought together Medicare, Medicaid, and commercial payers to provide financial and implementation support to the initiative. For example, in both MAPCP and CPC, a total of almost 40 payers participated in each initiative and in each state/region, where there were 4–9 payers. Medicare FFS participated in all initiatives except SIM and HCIA. Medicaid participated in CPC, MAPCP, and SIM. In Chapter 3 we describe some of the advantages and disadvantages of different payer structures.

2.2.2 Patient Populations Covered

Some of the initiatives also differed in terms of the populations they tried to serve. The IAH initiative had a narrowly defined patient population established by statute. Its population had the greatest number of disabilities: a majority of IAH beneficiaries required human assistance with five or six activities of daily living (ADLs) and had six or more chronic conditions. The population in the FQHC initiative was also unique. Because FQHCs serve vulnerable populations, the initiative served a high proportion of patients with disabilities; half of the Medicare patients in this initiative were receiving disability insurance, and about half were dually eligible for Medicare and Medicaid.

Exhibit 2-1. Overview of Characteristics of CMMI Primary Care Initiatives

	CPC	FQHC	IAH	MAPCP	SIM-Model Test Round 1	HCIA-PCR
Convener	CMS	CMS	CMS	State entity	State entity	Individual awardees
Participants	About 500 primary care practices operating in 7 geographical regions	About 500 FQHCs, mostly in rural areas, which provide primary care and serve 200+ Medicare beneficiaries	15 primary care practices, which provide home visits across 14 states to Medicare patients with chronic conditions/disabilities	About 850 practices in 8 states (VT, ME, NC, RI, NY, PA, MI, MN); practices include some FQHCs, rural health clinics, and CAHs	6 states (AR, MA, ME, MN, OR, VT); some overlap with MAPCP (ME, MN, VT) and CPC (AR, OR)	14 awardee organizations; only 8 included in this study* <i>(mix of community-based organizations, providers, payers)</i>
Participating Payers	Multi-payer model—commercial insurers, Medicare, Medicaid, CHIP(4–8 payers per region)	Single-payer model—Medicare FFS	Single-payer model—Medicare FFS	Multi-payer model, which varies by state—Medicare, Medicaid, and commercial insurers (4–9 payers per state)	Varies by state—Medicaid in all states; multi-payer in 3 states (AR, OR, and VT); Medicare is not participating in any state	Varies by award, but not Medicare or Medicaid
Period of Performance	October 2012–December 2016	November 2011–October 2014	June 2012–September 2017	July 2011–December 2016	October 2013–April 2018	June 2012–June 2015 (some June 2016)
Distinguishing Features	<ul style="list-style-type: none"> • Multi-payer model in 7 regions • Generous financial support, intensive TA, and data provision • Milestone approach to help practices transform 	<ul style="list-style-type: none"> • Single-payer • Low financial support from Medicare, although complemented by funding from HRSA and other sources • Goal was to have FQHCs become PCMHs and get NCQA recognition 	<ul style="list-style-type: none"> • Home-based care model, not PCMH • 15 experienced practices, so no up-front support and little TA from CMS • Patients had multiple chronic conditions and ADL limitations 	<ul style="list-style-type: none"> • Multi-payer model in 8 states • Variation by states in model, financial support, and TA • CMS joined ongoing initiatives in many states • Practices had to have PCMH recognition on entry or within 6–18 months 	<ul style="list-style-type: none"> • Goal is to transform the state’s health care system and have 80% of payments in each state under value-based or alternative payment models • Multiple interventions in each state • More emphasis on payment models and infrastructure (e.g., health IT/HIE) • Medicare FFS not a participating payer • 	<ul style="list-style-type: none"> • Each award was a separate delivery reform intervention, some were multiple interventions • Smaller initiatives, which were tailored to the specific local situations • Lump-sum awards • Little CMS involvement post-award •

Abbreviations: FFS, Fee-for-Service; CHIP, Children’s Health Insurance Program; CAH, Critical Access Hospital; ADL, activity of daily living *The HCIA awardees included in this study are (1) PeaceHealth, (2) Denver Health, (3) Finger Lakes Health Systems Agency, (4) Pacific Business Group on Health, (5) TransforMED, (6) CareFirst, (7) Cheyenne-WyPCMH, and (8) Sanford Health.

Exhibit 2-2. Initiative-Level Support to Practice

	CPC	FQHC	IAH	MAPCP	SIM-Model Test Round 1	HCIA-PCR
Payment Support	<p><i>Direct:</i> median payment of \$146,000 per practice per year from Medicare (including all payers, median payment of \$228,000 per practice)</p> <p><i>Indirect:</i> opportunity for shared savings</p>	<p><i>Direct:</i> average of \$31,000 per FQHC per year</p> <p><i>Indirect:</i> FQHCs also received financial support from HRSA in applying for NCQA PCMH recognition</p>	<p><i>Direct:</i> None</p> <p><i>Indirect:</i> opportunities for shared savings</p>	<p><i>Direct:</i> average of \$53,000 per practice per year from Medicare, plus payments from other payers; some direct funding for care coordinators/community health teams</p> <p><i>Indirect:</i> P4P in MI, NY, and RI; opportunity for shared savings in PA</p>	<p><i>Direct:</i> CMS funding to states, but no direct payments from Medicare to practices; some Medicaid payments to practices in AR, ME, and VT</p> <p><i>Indirect:</i> shared savings in 4 states</p>	<p><i>Direct:</i> CMS funding to awardees, but no direct payments from Medicare to practices</p> <p><i>Indirect:</i> P4P for CareFirst</p>
Technical Assistance	<p>CMS and its contractor team provided extensive TA through meetings, webinars, regional learning collaboratives, action groups and on-site coaching; a national prime contractor teamed with regional learning faculties (RLFs)</p>	<p>Multiple CMS contractors provided TA to help FQHCs achieve NCQA PCMH recognition, along with other transformation topics</p>	<p>Limited TA from CMS focused on meeting the initiative's administrative requirements and use of CMS data; the American Academy of Home Care Medicine administered a learning collaborative</p>	<p>States were responsible for TA and details varied by state; overall mix across the states included meetings, webinars, learning collaboratives, and one-on-one practice coaching</p>	<p>CMS provided TA to state staff; states provided a variety of TA to their practices</p>	<p>Most awardees provided internal TA for providers or implementation staff (e.g., for care coordinators); details varied by awardee</p>
Data Provision	<p>CMS provided quarterly practice-level feedback reports and patient-level data files, including cost, utilization, and quality measures; other payers also provided data; some regions made progress on alignment or aggregation across payers</p>	<p>CMS provided FQHCs with quarterly reports on readiness for PCMH accreditation as well as quality, cost, and utilization data for Medicare patients at the practice and beneficiary level</p>	<p>CMS provided quarterly practice-level and beneficiary level cost and utilization data for demonstration patients</p>	<p>CMS provided quarterly cost, utilization, and quality data to practices in five states; data separately provided by the states or other payers varied by state</p>	<p>CMS provided Medicare claims extracts to requesting state agencies; states varied on their data provision to practices</p>	<p>CMS provided Medicare claims extracts to requesting awardees; awardees varied on internal provision of data to practices</p>

The SIM initiative also served distinct populations that differed from the other initiatives. In SIM, four states are developing different types of medical homes to serve the special needs of medically complex Medicaid patients. In addition, four SIM states are developing behavioral health homes, which are PCMHs that are tailored for populations with mental health and substance abuse care needs. SIM Arkansas is also developing health homes tailored to Medicaid beneficiaries with developmental disabilities and those with long-term service and support needs. HCIA Sanford Health also focused on participants with mental health conditions.

2.2.3 Implementation of Interventions

The initiatives all shared the general goal of reducing or maintaining per capita health expenditures while maintaining or improving the quality of care. Some of the initiatives also included goals of improving access, coordination, and the patient experience, and some aimed to improve population health. There were differences in the interventions that were implemented. Some of these differences related to PCMH recognition, transformation goals, payment reform, health IT adoption and infrastructure, workforce development, and initiative milestones. We discuss each below:

- **Intervention goals.** Three of the initiatives (CPC, MAPCP, and FQHC) aimed to improve the quality of care by having practices (or clinics or FQHCs) adopt PCMH or other advanced primary care characteristics. The following are the six areas required to achieve the NCQA 2011 standards for PCMH recognition:
 - enhancing access and continuity
 - identifying and managing patient populations
 - planning and managing care
 - providing self-care support and community resources
 - tracking and coordinating care
 - measuring and improving performance

Some initiatives formally adopted NCQA PCMH recognition as a goal, while others allowed other types of PCMH recognition, and still others developed variants to the PCMH model standards (e.g., CPC established five key care delivery functions). In some initiatives, the conveners required additional activities not required by NCQA. Many of the MAPCP initiatives also required that practices be re-certified within three years.

SIM and HCIA were quite different from the other four initiatives in that they had many interventions. In SIM, each of the six states had multiple interventions. In HCIA, each of the eight awards was a different type of intervention (and in some cases, multiple interventions).

Some of the initiatives did not have primary care transformation as a goal. IAH practices were experienced at providing home-based care and did not expect to make significant changes in their health care delivery models. SIM attempted to transform each state's

overall health care system. As part of these efforts, SIM funds were used in some states to spread existing PCMH models to other payers, although no Medicare funds were paid directly to practices.

- **Payment reform.** SIM also differed in that it attempted to implement payment reforms. An explicit goal of the SIM initiative is to have 80 percent of provider payments in each state be subject to value-based arrangements or alternative payment models by the end of the initiative. Arkansas is testing episode-of-care models for both Medicaid and commercial payers.
- **Health IT adoption and infrastructure.** There were differences in the degree to which the initiatives attempted to improve health IT adoption, improve health information exchange, and build data infrastructure capacity. This was a key feature in some of the HCIA awards and in all six SIM states. Some of the MAPCP states also required special health IT efforts by practices. For example, MAPCP VT required practices to develop interoperability with a state clinical registry.
- **Workforce development.** SIM had a focus on workforce development focus in some states. For example, SIM Maine is piloting a project on integrating community health workers and SIM Oregon is investigating how to fill gaps in the care coordination workforce with traditional health workers.
- **Approach to achieving initiative goals.** There were also differences in the ways that the initiatives attempted to have practices achieve their goals. In CPC, specific annual milestones were established in nine areas that practices had to meet. Milestones were also established in the FQHC initiative, but not until well into the initiative's three-year period. In SIM and MAPCP, the states themselves decided on the goals and the ways that the initiatives would be monitored.

2.2.4 Practice Selection into the Initiatives

How sites were selected into the initiatives helps one understand whether the results of an initiative can be generalized to a broader group of practices and the degree to which initiative practices had already undergone primary care redesign. Although different approaches were used in each initiative, most of the initiatives established criteria that primary care practices would have to meet and then selected practices that met the criteria, as follows:

- For the CPC initiative, CMS initially selected seven regions and 31 payers to participate based on a payer application process. Interested primary care practices within each of the seven regions were then invited to apply. The CPC practice selection process favored practices that already had electronic health records (practices that met the EHR meaningful use requirements) and had previously engaged in transformation efforts, as measured by having PCMH recognition.
- For the FQHC initiative, CMS selected about 500 FQHCs from among the approximately 800 FQHCs that met CMS-specified eligibility criteria and that agreed to function as PCMHs. Only 7 percent of the FQHCs selected to be in the FQHC initiative had any type of PCMH recognition at the start of the demonstration, reflecting a low level of transformation.

- In IAH, CMS selected practices that met CMS criteria and that applied to be in the initiative. These practices had to have experience in delivering home-based care, although there was some variation with respect to their starting points, specifically regarding direct provision of, or coordination with, social work services and use of an electronic health record.
- For MAPCP, each state either selected the practices to participate in the demonstration or practices were grandfathered into the demonstration based on their participation in prior state delivery reform initiatives. For example, MAPCP Vermont was an extension of the state's Blueprint for Health initiative, which launched its initial PCMH pilot in 2007, well before CMS joined the multi-payer initiative. Similarly, practices already in existing programs in Minnesota, Michigan, Pennsylvania, and Rhode Island largely continued their participation, and additional state efforts expanded the multi-payer program to other practices within the state.
- The practices participating in the HCIA-PCR initiative were diverse and ranged from small independent practices to larger community health centers and hospitals. In all cases, the individual HCIA-PCR awardee organization selected the practices to participate, and for many this involved implementation of the proposed innovation among an existing network of practices or providers that the awardee owned or with which it had an existing affiliation. For example, in HCIA-TransforMED, a health IT intervention, the practices selected were based on their existing use of health IT and ability to accommodate new software. Practices in many of the HCIA-PCR awardees had exposure to PCMH concepts from previous projects the awardees implemented.
- In SIM, various approaches were taken to select practices in the primary care components. For example, Arkansas practices could become designated PCMH practices based on state-specific criteria and could opt in to participation during an open enrollment period. In Oregon and Vermont, practices were selected by virtue of their contractual relationships with participating payers. In Minnesota, state leaders implemented large integrated health partnerships (IHPs) serving the function of Medicaid ACOs; eligible entities applied to be recognized as an IHP through a request for proposal process from the state, and Section 2703 Health Care Homes (another component of SIM Minnesota) largely formed the practices within the IHPs.

A key difference between the initiatives concerns whether PCMH recognition was taken into account in practice selection. For CPC, CMS favored practices that were already PCMH-recognized as part of its selection criteria (about 40 percent of CPC practices had some type of PCMH accreditation at the start of the initiative). Although FQHC required that all participating centers achieve Level-3 NQCA 2011 standards for PCMH certification by the end of the initiative, at the beginning of the initiative only 7 percent of FQHCs had any NCQA PCMH certification. Five of the MAPCP states required PCMH recognition to enter the initiative and the other three MAPCP states required PCMH recognition within 6–18 months of entry. Because it did not pursue a PCMH model, IAH did not require formal PCMH recognition at any stage of the initiative.

In terms of generalizing the results of the initiatives to an expansion population, one must take into account that the practices were not randomly selected. The initiatives also included a

wide range of practices—from practices that had not transformed (FQHC) to practices that had started to transform or that were already PCMHs (CPC, MAPCP). Practices were selected based on their health IT/EHR capabilities in CPC and some of the HCIA awards. Because some of the HCIA practices were selected from practices that had previously been successful, the generalizability of the results from some of the HCIA awards is questionable. Another factor affecting generalizability of the initiative findings is that some of the initiatives focused on practices in rural areas. For example, three of the MAPCP states (VT, NC, and NY) primarily involved practices in rural areas.

2.2.5 Payment Model

Conveners decided the type of financial support that would be provided to practices: support could be provided *directly* to practices through per beneficiary per month (PBPM) payments for care management or *indirectly* through shared savings, pay-for performance (P4P), or bundled payments/episode-of-care payments. As shown in *Exhibit 2-2*, in many of the initiatives there was a mix of direct and indirect payments to the practices. Chapter 4 describes the advantages and disadvantages of different payment models in supporting practice transformation in more detail.

The three initiatives convened by CMS had conceptually simple payment models. In FQHC, direct per beneficiary per quarter (PBPQ) payments were made to participating FQHCs for beneficiaries who were attributed to the FQHC. The CPC initiative was more complicated in that not only were there direct PBPM care management payments made to practices for attributed beneficiaries, but practices were also eligible to participate in shared savings if regional savings targets were achieved and practice-level quality standards were met. Thus, the CPC payments both assisted in practice transformation (the direct PBPM payments) and in providing an incentive for practices to be efficient and high-quality (shared savings). The third CMS-convened initiative, IAH, provided no direct payments to practices in accordance with its enabling legislation but allowed shared savings to be earned if practices met both quality standards and achieved savings targets.

MAPCP, FQHC, and CPC provided direct PBPM payments to practices. However, many MAPCP states also paid *non-practice* entities to provide care coordination, analytic support, or other services to participating practices and their respective patients. The types of entities included community care or community health teams, clinical support vendors, physician support organizations, and analytic support/health information technology consultants. For example, four MAPCP states (ME, MN, NC, VT) made payments to community care or community health teams, and two additional MAPCP states (MI and NY) made payments to physician associations for the provision of care management/coordination services. The MAPCP payments to non-practice entities were substantial. In MAPCP NC, for example, a PMPM of \$2.50-\$3.50 was paid to each practice, and a \$6.50 PMPM was paid to a community care network.

MAPCP was similar to CPC in offering both PBPM and shared savings in one state (PA). In addition, three MAPCP states had both PBPMs and P4P provisions (MI, NY, and RI). A distinction between the CPC and MAPCP payment models is that the payment methods in MAPCP had to be applied consistently by all payers in a state, although the payment levels could

differ. The SIM and HCIA initiatives differed from the other four initiatives. In SIM, Medicare made no direct payments to practices. One SIM state, Arkansas, developed bundled payments for episodes of care to encourage efficiency and to transform practices away from fee-for-service incentives. Much of the SIM funding was used for infrastructure investments in building health IT and data analysis capabilities and developing learning collaboratives. The payment models used by HCIA-PCR awards largely reflected existing payment schemes in place at the awardee organization and its partners.

2.2.6 Beneficiary Attribution

Beneficiary attribution is the method used by conveners to determine whether a practice receives a PBPM payment. The methods used by the initiatives varied, but all attribution methods used information about a Medicare beneficiary's utilization (typically the proportion of evaluation and management [E&M] services) to determine whether he or she should be assigned (i.e., attributed) to a practice. Most initiatives did the assignment retrospectively, which means that practices often did not know whether they were responsible for the primary care of a beneficiary until after the fact. Some initiatives enrolled beneficiaries (e.g., IAH and MAPCP-MN).

2.2.7 Technical Assistance

Participating practices across most of the initiatives received some form of technical assistance (TA) to help them move toward practice transformation. *Exhibit 2-2* summarizes the technical assistance provided across the various initiatives; Chapter 4 discusses lessons learned about TA across the initiatives.

TA topics offered in most initiative settings included training in care coordination, practice-level quality improvement, implementation and use of health IT, use of practice feedback reports, and meeting initiatives' administrative requirements. Several initiatives also offered TA to help practices attain PCMH recognition (FQHC, five MAPCP states, HCIA-Cheyenne, and SIM AR and OR). Several initiatives also offered TA related to staff development or the integration of new professions into care delivery teams. The IAH initiative differed sharply from the others in that the TA provided by the convener (CMS in this case) was limited to training on the initiative's administrative requirements, such as data reporting, and explanation of CMS-provided data reports. However, the IAH practices did start their own learning collaborative supported by an external organization (the American Academy of Home Care Medicine).

Initiatives used multiple modes for delivering TA. Many initiatives provided individualized, one-on-one assistance to practices, such as on-site or virtual practice coaching. A majority of initiative settings also conducted learning collaboratives for peer-to-peer interaction using a mix of face-to-face meetings and virtual exchanges. Webinars were another common training mode, while a few cases (e.g., FQHC, MAPCP MN, and MAPCP NC) also developed online toolkits for practices. Some initiatives also provided more formal, classroom-style training of practice staff for topics such as use of intervention-specific technology, care coordination, and patient education.

2.2.8 Practice-Level Data Support

In addition to the learning activities discussed above, these initiatives also typically provided data to participants, as summarized in *Exhibit 2-2*. Chapter 4 discusses lessons learned about data provision across the initiatives. Initiatives provided cost, utilization, or quality data in CPC, FQHC, IAH, all eight MAPCP states, most SIM states, and five of the eight HCIA PCR awardees included in this study (CareFirst, Finger Lakes, PBGH, PeaceHealth, and TransforMED). In most cases, the data included practice-level reports and patient-level reports but not uniformly. Most initiatives provided these data quarterly, although the substantive content reflected longer time lags from the date of care due to the time involved in processing claims and producing the reports.

The capacity of practices to access and use these data varied widely, primarily because of differences in staff capacity and health IT challenges. Most initiatives also provided training related to use of these data reports, but the degree and level of training varied widely.

Some initiatives offered additional data infrastructure to practices, such as all-payer claims databases, health information exchanges, or support in obtaining near real-time alerts when a practice's patient was discharged from an inpatient hospital stay or an emergency department visit.

2.3 Conclusion

The initiatives differed quite substantially on many dimensions, even though all hoped to improve quality while reducing costs. Some of the distinguishing features are summarized in *Exhibit 2-1*. The CPC and MAPCP initiatives are the most similar: each used a multi-payer model, a per member per month (PMPM) payment model (though payment levels were more generous in CPC and many MAPCP states authorized payments directly to community care teams for care coordination), and offered similar kinds of TA and data feedback to practices. A key difference is that CMS convened CPC, while the states convened MAPCP.

FQHC is similar to MAPCP and CPC with its clear focus on office-based primary care practice transformation to PCMH; however, this CMS-convened initiative is single-payer (Medicare) and provided much lower payments to practices for transformation. Furthermore, the beneficiaries using FQHCs were more likely to be low-income.

IAH is unlike any other initiative in terms of its goals and focus on using a home-based primary care model to care for highly disabled chronically ill beneficiaries. IAH is also the only initiative that did not include any requirements or assistance with respect to formal PCMH recognition. In addition, the IAH practices received no up-front financial support and little TA.

The HCIA-PCR initiatives also differ from the other initiatives in that they are not convened by CMS or the states. They focus exclusively on delivery innovation, as opposed to payment innovations, and no two cases are alike, though many share a common focus on offering care management and coordination services.

Finally, the SIM states are largely building on the existing primary care infrastructure that already existed in those states. Most SIM states involve several components beyond primary care

redesign and practice transformation. Some SIM states do include efforts to establish Section 2703 Health Homes and Behavioral Health Homes, but these are generally focused on the Medicaid population.

3. What are the Challenges and Opportunities in Convening Primary Care Initiatives?

This chapter addresses challenges that conveners encountered in designing and implementing the initiatives and what lessons were learned about addressing these challenges.

3.1 Convener

Key Findings on Conveners

- Transformation initiatives that are largely run by the states rather than CMS take advantage of local relationships and allow tailoring for local circumstances, but implementation progress can be slow because states tend to place a high priority on within-state relationships and generally have limited and sometimes fluctuating resources in terms of staff and expertise.
- In designing initiatives, it is important to define the roles and responsibilities of all players. Successful collaboration among multiple stakeholders requires conveners to clearly communicate the roles and responsibilities of the various payers and providers as well as the goals and strategies of CMS.

The convener is the entity responsible for leading an initiative or setting and organizing its activities. Conveners include state agencies, CMS, commercial payers, and health care systems. The conveners decided how each initiative was designed, how primary care transformation was to take place, and which practices were included in each initiative.

Selecting practices experienced in implementing transformation activities likely facilitates initiative implementation. According to initiative evaluators, several practice characteristics were associated with the ease with which practices implemented primary care transformation. For example, key informants from four of the six initiatives (CPC, FQHC, HCIA-PCR, and IAH) noted that practices with a lack of experience with primary care transformation faced unanticipated challenges in that they did not fully understand what primary care redesign entailed, needed additional technical assistance, set overly ambitious goals, and took longer to become operational. In addition, these practices were often unprepared for the requirements, terms, and conditions of a CMS initiative. In addition, transformation was harder for small practices with fewer resources.

3.1.1 Different Origins of Initiatives

There were benefits and challenges to both state-convened and CMS-convened approaches. In this section, we review the structure, benefits, and challenges of different convener and stakeholder arrangements and how these challenges were overcome (*Exhibit 3-1*). CPC and FQHC were initiatives developed by CMS staff by issuing requests for proposals from private payers and practices to voluntarily participate. In contrast, CMS joined existing state-sponsored multi-payer initiatives for the MAPCP and SIM initiatives. IAH selected 13 practices and one consortia to participate, and HCIA-PCR included competitive awards to health care delivery systems and payers that, respectively, met multiple criteria for inclusion (largely stipulated by Congress) or proposed innovative ways to improve the quality and lower the cost of care for Medicare, Medicaid, and Children's Health Insurance Program (CHIP) enrollees.

Exhibit 3-1. Benefits and Challenges by Type of Convener

Convener	Benefits	Challenges
CMS	<ul style="list-style-type: none"> • Contributes financial and technical resources • Staff capacity for long-term management • Can use policy levers to encourage participation and facilitate implementation • Can use purchasing power to encourage participation • Flexibility to make decisions 	<ul style="list-style-type: none"> • Playing dual roles of convener and collaborator • Lack of familiarity with local players and context
State Agency	<ul style="list-style-type: none"> • Can use purchasing power to encourage participation • Familiarity with local players and context • Long-term relationships with stakeholders • Can build on existing programs (if any) 	<ul style="list-style-type: none"> • Competing priorities makes decision-making more difficult • Turnover among state-level management • Siloed state structures hinder collaboration around implementation
Other (i.e., payers, health systems)	<ul style="list-style-type: none"> • Familiarity with local players and context • Can build on existing programs (if any) • Fewer regulatory hurdles 	<ul style="list-style-type: none"> • No legislative or regulatory power to facilitate implementation • Health systems have no purchasing power to encourage participation

Regardless of the convener, our survey of evaluators conducted in the spring of 2016 revealed broad leadership teams in all initiatives with conveners, leaders, payers, and other stakeholders collaborating on initiative design and implementation. Most (23 of 31 settings)³ leadership teams involved state agencies or state-level entities, either as the sole leader or part of a coalition. For example, the Michigan Department of Health and Human Services convened MAPCP-MI and included a steering committee of payers in its leadership team. The CPC settings were led by coalitions in each region composed of CMS staff, participating payers, TMF Health Quality Institute (the prime learning contractor across all CPC regions), and stakeholder groups.

Even though CMS was not the convener in MAPCP and SIM, CMS involvement helped state-run initiatives by encouraging payer participation and providing financial and in-kind supports. With MAPCP, for example, CMS provided practices with monthly fees,

³ The 31 settings refer to the seven CPC regions, eight MAPCP and six SIM states, eight HCIA-PCR awardees, and the FQHC and IAH initiatives.

practice feedback reports, and training on the web data portal. When debates arose in SIM-AR about the definition of population health, CMS provided a template for the convener and stakeholders to work from. Generally, key informants indicated that initiatives valued CMS's participation for encouraging payers to participate and for contributing financial and technical resources.

CMS faced challenges in playing the dual roles of convener and collaborator in CPC. Payers generally valued CMS's active participation in CPC and indicated that CMS involvement encouraged broad payer and practice participation and brought financial and technical support to their regions (Taylor et al., 2015). However, some payers expected more collaboration with CMS and did not anticipate all of the requirements for payer meetings, performance measurement, data aggregation, and learning activities. For example, "payers indicated that CMS initially signaled that regional payers would collaboratively select common measures to track, but later clarified that the region could add only a few measures to those already selected by CMS for the national initiative" (Taylor et al., 2015, p. 77).

Evaluators in state-convened initiatives thought that the state's often long-term relationships with stakeholders, their familiarity with local players and context, and capacity to build on existing relationships and programs was a benefit. In spite of these benefits, states typically have more limited resources and competing priorities—a situation exacerbated when siloed state structures hinder collaboration. When state structures are siloed, they may have competing priorities, compete for resources, and have restricted lines of communication across agencies.

Other conveners (e.g., payers, health systems) also benefit from long-term relations and face fewer regulatory hurdles but have no legislative or regulatory power to facilitate implementation and health systems have no purchasing power to encourage participation.

3.1.2 Initiative Goals, Roles, and Responsibilities

In a single-payer initiative, the convener is able to design the goals of an initiative to meet its objectives, while multi-stakeholder teams require clarity on goals, roles, and responsibilities to maintain a productive and harmonious collaboration. Stakeholders sometimes lacked clarity on the parameters of the initiatives. For example, stakeholders of SIM-MA reported that they understood the individual activities with which they were involved but did not understand the full scope of SIM-funded activities in the state (Gavin et al., 2016). In response, the convener planned to reset and communicate its vision for the initiative. Likewise, stakeholders involved with the SIM and CPC initiatives wanted more detail on roles and responsibilities. Consumer representatives involved with CPC-NY and CPC-OH/KY wanted to know specific ways they could contribute.

Communication challenges can emerge when there are multiple conveners or leaders of initiatives. Payers asked for greater clarity and consistency on CMS's involvement. Furthermore, payers thought that CPC could "be improved if CMS capitalized on local experience and priorities" (Taylor et al., 2015, p. 77). Payers said that CMS should have consulted them about strategy and the practices selected so that they could have shared their expertise on their regions and which practices were high and low performers. It was felt that greater communication, role

clarity, and capitalization on payers' regional expertise would have helped build more collaboration.

As clarity on goals, roles, and responsibilities improved, so too did satisfaction with the initiatives and their involvement. For example, members of SIM-MN's task force reported that they were uncertain of their roles in the beginning, but over time the group matured and became more cohesive and productive with clear goals and actionable discussions. They accomplished that by giving task force members homework assignments to provide feedback to the conveners via email, and by focusing the meetings on specific topics.

3.2 Multi-Payer Structure

Key Findings on Multi-payer Structure

- Multi-payer structures help practices transform because they increase the financial resources to practices and the number of patients in a practice that use modified care protocols.
- The benefits of the multi-payer structure are greatly increased if the payers can agree upon a set of harmonized quality measures and their reporting.

A key design decision faced by conveners is whether to have a multi-payer structure, in which multiple payers participate in and provide financial support to a primary care initiative. Because CMS has participated in both single-payer and multi-payer structures in these six primary care redesign initiatives, we have been able to identify some of the advantages and disadvantages related to these payment structures.

It is much simpler for CMS to design a single-payer initiative because the convener is able to make design decisions without getting the approval of other payers. As a result, these initiatives can be implemented more quickly than under a multi-payer structure. Single payer also allows CMS to implement an initiative nationally, if desired, because it is not necessary to focus on locations with other willing payers. However, a major drawback to an initiative convened by a single payer is that the payer is typically able to provide only limited financial support to an individual practice because it only makes financial contributions for its share of a practice's patients. This can lead to a low overall level of financial support, especially if the payer accounts for a small percentage of patients in a practice, as was the case in the FQHC initiative where, on average, the Medicare patients accounted for only 8 percent of patients.

Multi-payer initiatives allow more financial support and an opportunity to align transformation requirements or activities across most or all of a practice's patients. A more fundamental challenge of single-payer structures, which is related to the level of financial resources that a single payer can contribute, is whether a single-payer initiative can influence how practices transform care for their patients. Single-payer initiatives fund a practice for one group of patients, not the entirety or necessarily even the majority of the practice's patients. Thus, it is much more difficult for a single-payer initiative to impact the way that an entire practice delivers care. However, the ability of single-payer initiatives to substantially impact practice transformation varies for two types of practice transformation activities:

- Some practice-wide changes, like modifications to the practice's EHR, have benefits that are likely to spill over to all of the practice's patients. In situations like this, if a practice has sufficient financial support (even from a single payer) to purchase or modify an EHR or extend its office hours, all of the practice's patients will get the benefits. However, a single-payer initiative is less likely to provide the level of financial support necessary to make this type of investment.
- A second type of practice transformation activity typically does not result in much spillover to patients not covered by the initiative. For example, new staff hired to work as care coordinators often focus exclusively on patients of participating payers. This means that in single-payer initiatives, practices may only be financially supported to modify procedures for that group of patients. In addition, practices often are burdened by having to identify patients who are eligible for these enhanced services and those who are not. This makes it more complicated to modify the practice's routines and makes practice transformation less likely.

There are, however, design challenges with multi-payer structures as well. One challenge is that each payer in a multi-payer structure may establish different requirements for its patients regarding new care management protocols, quality measures, and reporting. In the MAPCP initiative, one informant reported that a single patient in a multi-payer initiative had three different care managers. Not only is this an operational issue for practices, but it also increases the administrative burden and costs for practices if they must follow different sets of payer rules for different patients (just as they do in single-payer initiatives). In some initiatives, this challenge was overcome by initiative leaders agreeing upon a common set of reporting requirements and quality measures. Practices expressed a strong interest in harmonizing quality measures and their reporting, although harmonizing measures can be difficult and time-consuming because it can be contentious to get some private insurers to agree upon a common set of measures. Many payers regard their approaches to quality measurement and other aspects of value-based purchasing as a unique, proprietary tool and are reluctant to reduce their competitive advantage by agreeing to use the same set of payments and reporting requirements as other payers. The CPC initiative addressed the potential problem of having different approaches apply to different patients by requiring that all changes made as part of CPC be delivered to all patients in the practice (Peikes, Anglin, et al., 2016). The efforts in having multiple payers meet to harmonize measures also allowed payers to share lessons they had learned, which is another benefit of a multi-payer structure.

Another design challenge in multi-payer structures is that the recruitment of multiple payers is time-intensive. This challenge was overcome in the MAPCP initiative when CMS joined ongoing efforts of states that had already recruited payers. Key informants reported that CMS joining an initiative was viewed as a positive sign by potential payer participants and encouraged other payers to join in transforming primary care.

Despite these challenges with a multi-payer structure, key informants thought that multi-payer structures were preferable to single-payer structures. Key informants associated with multi-payer initiatives thought that the multi-payer structure was essential to the prospects for success of their initiative, while key informants in single-payer initiatives thought that their initiative could have been improved by involving other payers. To reduce the burden of transformation on practices in a multi-payer initiative, it is important for payers to collaborate in

designing the initiative and harmonizing its rules, measures, and incentives. In some of the MAPCP initiatives where payers had already collaborated in previous efforts, they were able to build upon existing relationships among stakeholders.

3.3 External Contextual Factors

Key Findings on External Contextual Factors

- Statewide initiatives can be affected by changes in state political offices, especially shifts in the governor's office. Limited staff resources at the state level means that implementation of statewide primary care initiatives may compete with implementation of other state priorities.
- Competing initiatives, including participation in ACOs, made complying with initiative requirements difficult and caused some practices to not participate in or drop out of the primary care initiative.

External contextual factors influence initiative implementation. Three main external contextual factors affected implementation: (1) local and state policies and politics, (2) overlapping health care reforms, and (3) market characteristics.

3.3.1 Local and State Policies and Politics

State politics presented challenges to implementing statewide initiatives. Six of the eight MAPCP states elected new governors during the initiative. In some cases, newly elected state executives and legislators had different priorities than previous elected officials and the initiatives became lower priorities. For example, a new Pennsylvania administration dissolved the office in charge of the MAPCP initiative, which caused MAPCP implementation staff to resign. The governor also lifted the mandate requiring Medicaid Managed Care Organizations to participate. As a result, some payers discontinued their participation in the initiative, which led to some practices abandoning the initiative (McCall et al., 2015a).

In another example, SIM-MN stakeholders lobbied for the state legislature to exempt participating practices from strict HIPAA regulations that made sharing personal health information difficult, but the legislature did not approve the requested exemption (Gavin et al., 2014). This made coordinating care across collaborating partnerships difficult.

Establishing health insurance exchanges under the ACA created competing priorities in some SIM states, although in other states the ACA provided additional resources necessary for practice transformation. For example, because state-level SIM data analysts had to first address challenges from the launch of Oregon's ACA Health Care Exchange, their participation in implementing the SIM initiative was delayed (Gavin et al., 2014). On the other hand, FQHCs benefited from ACA grants that provided support for practice transformation. These up-front resources allowed under-resourced FQHCs to develop infrastructure critical for transforming the practices, such as expanding services and access to care. The FQHC evaluation found participating FQHCs that also receive funding from ACA grants were more likely to achieve NCQA Level-3 recognition (Kahn et al., 2016).

3.3.2 Overlapping Health Care Reforms

Practices facing competing priorities (most often other health care reform efforts) had difficulty complying with initiative requirements. For example, many FQHCs had multiple sources of funding and had to meet multiple reporting requirements. Private funders required additional measures and added reporting requirements but provided little additional money to support those efforts (Kahn et al., 2016). As one FQHC key informant explained, if a conflict among requirements arose, compliance with the FQHC initiative requirements tended to suffer. However, CMMI staff attempted to mitigate this challenge by aligning timelines with competing initiatives. Even when health care reforms supported FQHC transformation efforts, they required documentation, and the range of other quality and safety programs, requirements, and improvement initiatives competed with the PCMH effort for attention and resources. For FQHCs, these competing priorities included the following (Kahn et al., 2016):

- EHR implementation and meaningful use requirements
- Implementation of the International Classification of Diseases, Tenth Revision (ICD–10) implementation
- Private payer quality and safety initiatives
- HRSA requirements and initiatives, such as Uniform Data System (UDS) measurement reporting and performance standards

The documentation required by the various reform efforts strained practice capacity. As one FQHC respondent put it, “There are so many documentation requirements ... [and] being an FQHC, we don’t have all that staffing, ancillary staff, to really assist the providers in doing all the documentation” (Kahn et al., 2016, p. 76). Other practices participating in multiple health care reform efforts also found meeting requirements of the primary care initiative difficult. One HCIA key informant described competing initiatives as an issue across some HCIA awardees, particularly the HCIA awardees also involved with the FQHC initiative or other regional activities. Concurrent primary care initiatives also caused problems among participating practices. Each MAPCP state had other single-payer practice transformation initiatives in progress as well as federal initiatives, such as SIM and ACOs. Three of the MAPCP states were also SIM states. Although these practice transformation efforts can complement each other, competing initiatives stress the capacity of practices to meet initiative goals and requirements (Nichols et al., 2016).

ACOs decreased practice participation in primary care initiatives. Practices saw ACOs as an alternative to participating in traditional primary care initiatives, although some initiatives prohibited concurrent participation in ACOs. Of the 502 CPC practices, 29 (5.8 percent) left the initiative to participate in an ACO, and some payers reported that some ACOs encouraged practices to drop out of the CPC initiative and join the ACO (Peikes, Ghosh, et al., 2016; Peikes, Anglin, et al., 2016). Some practices also left the MAPCP initiative in Pennsylvania because the practices were already part of an ACO. (The eligibility requirements in Pennsylvania precluded practices in an ACO from participating in MAPCP.)

Initiatives running concurrently can also bolster each initiative’s implementation. Concurrent initiatives were mutually supporting when requirements aligned or one initiative built

the foundation for transformation efforts of another initiative. For example, as part of the SIM initiative, the Maine Health Management Coalition produced practice feedback reports for primary care providers (PCPs) across the state of Maine. PCPs participating in SIM, MAPCP, and other initiatives used feedback reports to inform practice transformation efforts (Nichols et al., 2016). In fact, the states that designed the SIM initiative were able to build off of existing practice transformation initiatives in all six states.

3.3.3 Market Characteristics

Practices in states or regions with experience with practice transformation and infrastructure to support practice transformation advanced faster than those practices in states or regions without prior experience or infrastructure. Existing regional infrastructure, a history of collaboration with partners, and legislative mandates (for example, an EHR mandate in Minnesota) facilitated practice transformation. Regional infrastructure includes a health information exchange (HIE), learning collaboratives, and regional culture that is ready for change. As one CPC key informant reported, practices in regions with existing infrastructure and in states that implemented statewide primary care initiatives prior to the CPC initiative more easily implemented primary care transformation as compared to practices in states without such initiatives. Where PCMH and practice transformation was a priority prior to the primary care initiative, practices progressed faster in meeting PCMH requirements relative to practices in states where PCMH and practice transformation were introduced through these initiatives.

Changes in and the complexity of the health care markets affected practice participation. Changes included mergers, acquisitions, and payment model reform. Mergers and acquisitions discouraged practice participation, whereas payment model reform tended to encourage it. For example, in North Carolina, there were many practice mergers and acquisitions of practices by larger practices or health care systems. In the case of an acquisition, participation in MAPCP was determined by the larger health care system, not by the practice being acquired. In addition, newly merged practices did not always have the required PCMH recognition to participate in the MAPCP initiative.

Alternatively, payment model reform enhanced practice participation by facilitating practice transformation through additional funding support. For example, the transition from fee-for-service to value-based purchasing can provide new financial support for many practice transformation activities, although how value-based purchasing is implemented likely affects transformation success (Nielsen, Buelt, Patel, & Nichols, 2016). Some Medicaid agencies now reimburse for community health worker services and telehealth services. Other payers provide additional payments to practices to support PCMH recognition. Lastly, by creating or participating in an ACO, practices have received support through capitated or shared savings from Medicare. These payments align with many initiative goals and facilitate practice transformation activities by reinforcing funding for these activities (Gilman et al., 2014).

4. What Facilitates Primary Care Transformation at the Practice Level?

The experiences in the six initiatives allow one to examine the challenges faced by practices in changing the ways in which they delivered primary care. This chapter describes some of these challenges and the types of support that helped facilitate practice transformation. Section 4.1 describes some of the ways that initiatives tried to support practices. Section 4.2 describe the features of the practices themselves that were most important in facilitating practice transformation and overcoming its challenges.

The chapter addresses the challenges that practices faced in the initiatives, how these challenges were overcome, and which initiative features best supported transformation.

4.1 Initiative-Level Facilitators

4.1.1 Financial Support

Key Findings on Financial Support

- Practices need different levels of support at different levels of transformation. Up-front payments or higher payments at the beginning of the transformation may be necessary.
- To pay for the costs of care managers, small practices may need higher PBPM rates than larger practices.
- Shared savings provisions raise multiple issues, especially for small practices, but some of these challenges can be addressed by pairing shared savings with PBPM payments.

The level and type of financial support practices need may depend in part on practice's current level of readiness for, or progress toward, practice transformation. Practices in the initiatives needed financial support because Medicare's fee-for-service payments do not reimburse practices for some important activities required by advanced primary care practices, such as care management, care coordination, and reporting. The level of financial support required for practice transformation depends on a number of factors, including the following:

- The current level of transformation of the practice (e.g., whether it already has an EHR, is a recognized PCMH, or whether it employs a care manager)
- The requirements of the initiative (e.g., what type of PCMH recognition is required)
- Other resources that a practice has access to, including from other payers, in order to pay for transformation activities

There is no agreed-upon payment strategy that has been proven to be most successful in primary care redesign (Nielsen et al., 2016). Nor is there agreement on the required level of financial support (Magill et al., 2015).⁴ One reason is that designing the level and form of financial support that will best support practice transformation is challenging because one model of reimbursement may not be appropriate for every phase of medical home transformation (Barr, 2010). For example, practices without well-functioning EHRs can require relatively large, up-front expenditures for health IT. Moreover, even practices with adequate EHRs may still require

⁴Michael Magill and colleagues estimate the costs per full-time equivalent primary care physician associated with PCMH functions to be in the range of \$7,700 to \$9,700 per month.

care managers and other additional staff to become advanced primary care practices, which requires adding costly personnel like nurses. The six initiatives used different approaches to provide financial support to practices and allow us to assess what their effects have been.

4.1.1.1 Levels of Financial Support

Financial support was provided either directly to practices through care management fees (typically PBPM payments) or indirectly through such mechanisms as shared savings or pay-for-performance (P4P) incentives. In three of the initiatives (FQHC, CPC, and MAPCP), Medicare made PBPM payments to practices, although some of the Medicare PBPM payments in MAPCP went to non-practice entities that, in turn, provided care coordinators to practices. In these three initiatives, the PBPMs were the primary source of financial support. In the initiatives with shared savings or P4P incentives (CPC and MAPCP), the aggregate level of PBPM payments greatly exceeded the value of incentive payments.

There are alternative ways to compare the level of financial support across initiatives. One approach is to focus on the PBPM that is paid for each attributed beneficiary. As shown in *Exhibit 4-1*, the Medicare PBPMs varied from \$6 in FQHC to an average of \$20 in CPC (after the first two years, the CPC average PBPM dropped to \$15). Direct Medicare support to MAPCP practices averaged \$5.47 PBPM, but when Medicare payments to non-practice entities like community health teams are included, the Medicare payments averaged \$8.90 PBPM. A second approach is to compare the level of financial support provided to the practice. The actual Medicare payments per practice, which represent only a portion of the available financial support of initiatives that included multi-payer participation, varied more than the PBPMs. In CPC, the median practice received almost five times as much in Medicare payments as the median FQHC, not only because the PBPM amounts were higher in CPC, but also because the CPC practices had more Medicare patients. From the practice's perspective, the overall level of financial support also varies by whether an initiative is multi-payer or single-payer. In multi-payer initiatives like CPC and MAPCP, Medicare's payments were supplemented by payments from Medicaid and private payers. In CPC, this increased the median payment per practice in 2013 from \$146,000 (Medicare alone) to \$228,000 when considering all payers. Because practice sizes vary, the payments per practice or PBPM may not be the best indicator of support to clinicians in the practices. As a result, we calculated a third measure of financial support: the level of Medicare support per clinician. Although we were only able to calculate the annual level of Medicare support per clinician for CPC and MAPCP, we found a wide range of payments per clinician: the average annual Medicare payment per clinician was approximately \$34,900 in CPC and \$10,500 in MAPCP (excluding MAPCP-MN).

4.1.1.2 Use of Financial Support

Particularly for small practices, up-front payments or higher payments at the beginning of the transformation process may be necessary to pay for the costs of hiring care management staff, upgrading practice health IT capabilities, and changing treatment protocols. The CPC evaluation included a thorough analysis of how the financial support was used by practices. This evaluation found that 90 percent of CPC PBPM payments were used for labor costs, which was defined to include the cost of hiring physicians, care managers, RNs, and other labor (Peikes, Anglin, et al., 2016). Practice surveys from the FQHC initiative indicated that the PBPMs were used to hire staff, modify EHRs, conduct staff education and training, and develop patient

information materials, although the evaluation did not quantify the distribution of payments among these categories. The FQHC evaluation also found that about half of the FQHC site respondents were unaware of how the FQHCs spent their funds; similarly, the evaluation of the MAPCP initiative found that many practices were unable to say how the PBPMs funds were used.

Exhibit 4-1. Annual Medicare PBPM Payments to Practices in Three Initiatives

Initiative	Average Level of PBPM	Annual Medicare Payment Per Practice	Multi-Payer?
CPC*	\$20/\$15	\$146,000	Yes
MAPCP**	\$8.90/\$5.47	\$57,800	Yes
FQHC***	\$6	\$31,400	No

*CPC PBPM varied by risk level (HCC). CPC payment per practice is the median amount in 2013 (CPC, ARI, p.46). CPC PBPM dropped from \$20 to \$15 after the first two years of the initiative. **MAPCP PBPMs varied by state (MAPCP, FR, p.3-93). MAPCP PBPM value shown is a weighted average payment to practices, excluding Minnesota, across all years of the initiative. The higher PBPM (\$8.90) is the value of Medicare payments to practices and to other entities, such as community health teams. The lower PBPM value (\$5.47) includes only the payments to practices. ***FQHC payment per practice is the median across the initiative. (FQHC, FR, p.23).

Note: The SIM, HCIA, and IAH initiatives were excluded because there were no Medicare PBPM payments.

4.1.1.3 Adequacy of Support

A key challenge to practice transformation is whether the practice receives sufficient financial support to pay the additional costs required by transformation. We examined the adequacy of financial support from two perspectives. First, for the three initiatives that paid PBPMs to practices, we reviewed the results of the practice surveys and interviews conducted by the evaluators:

- Interviews with site leaders at the FQHC demonstration sites found that the amount of funding provided by the single-payer FQHC demonstration was “generally insufficient to cover PCMH implementation” (Kahn et al., 2016, p. 24).
- The survey of CPC practices in 2014 found that 81 percent of surveyed practices thought that the average Medicare PBPM of \$20 was “adequate” or “more than adequate” relative to the costs of implementing CPC. In program Year 3 (2015), the average CPC PBPM for Medicare patients was reduced to \$15. In that year, 76 percent of the surveyed CPC practices thought that this lower Medicare payment was adequate or more than adequate relative to the costs of implementing CPC. When the PBPMs were reduced in 2015, “there was no evidence of practices reducing staffing, and no complaints from practices about reductions in fees” (Peikes, Anglin, et al., 2016, p. 29). The CPC evaluators did a “deep-dive” analysis of CPC practices and

found that the level of CPC funding was a particular problem for very small practices (e.g., 1–2 physicians), which struggled to pay for a full-time care manager.

- The MAPCP evaluation found that in most states practices said the Medicare payments were insufficient to cover the costs of *all* the practice enhancements, though Michigan was the one exception. We calculated that the Medicare payment per practice in Michigan averaged almost \$70,000 per year per practice, which was about 50 percent higher than the average annual Medicare payment per MAPCP practice in the other states.⁵

A second perspective on the adequacy of financial support is to examine whether the financial support allowed practices to meet the initiative requirements. Many FQHCs received a HRSA grant that covered the cost of applying for NCQA recognition. With the additional median payment of approximately \$31,000 per year from Medicare, about 70 percent of the FQHCs in the initiative were able to achieve NCQA Level-3 PCMH recognition by the end of the initiative.

Meeting the initiative requirements is not equal to practice transformation. For example, one conclusion from the FQHC evaluation was that the enhanced payments during the demonstration were not enough to sustain the changes that occurred during the demonstration and that NCQA Level-3 PCMH recognition was not the same as transformation to a PCMH.

4.1.1.4 Shared Savings

Shared savings incentives raise multiple challenges for practices. One method of designing affordable initiatives has been to reduce the cost of an initiative through the use of shared saving provisions. Shared savings typically provide payments to practices that meet a specified savings target while meeting specified quality standards. Shared savings also provide an incentive for performance. One recent review of evidence concluded that PCMH initiatives that had the most impressive cost and utilization results were those that participated in multi-payer collaboratives with specific incentives or performance measures linked to quality or cost savings (Nielsen et al., 2016). Another important advantage of this type of design is that the government does not make payments to practices unless there are savings. However, numerous challenges exist with using this payment design feature to support the costs of practice transformation.

Two problems with shared savings are their timing and uncertainty. Savings cannot be measured until well after a performance year is over, assuming savings are evaluated on an annual basis to avoid the volatility of quarterly data. In the meantime, practices need to hire care management staff, upgrade their health IT, and change treatment protocols. This involves substantial financial commitment given the uncertainty of achieving rigorous savings targets. This additional commitment, in turn, can make recruitment or retention of practices more difficult in voluntary initiatives, especially for smaller practices that are less able or willing to bear risk.

A third challenge in using shared savings payments to support practice transformation costs was that the process was often contentious and time-consuming. For example, participating

⁵ Due to billing problems, Minnesota was an outlier and was excluded from this calculation.

practices complained that lack of savings was based on a flawed comparison group. Additionally, some practices criticized the list and measurement of quality outcome measures used as a prerequisite for shared savings payments.

Another challenge to using shared savings is not so much their direct impact on supporting practices but their effectiveness as an incentive. The strength of these incentives depends on how savings are measured and distributed to practices. For example, in the CPC initiative, savings were measured, collectively, across all participating practices in a region rather than at the level of individual practices that are directly bearing the costs of transformation. Any asymmetric distribution of costs and savings creates a *free rider* problem, and key informants thought this regional measurement degraded the incentive. Inefficient practices might share in savings without changing practice patterns at all. This challenge to shared savings may be extreme if practices are part of a larger system of practices that chooses not to distribute shared savings paid to the payment system to individual practices.

To overcome some of the challenges in supporting practice transformation costs and incentivizing increased quality and reduced costs, some initiatives combined shared savings with PBPM payments. This approach has the advantage of providing up-front, non-refundable payments (the PBPM) to practices for capital and labor investments while still providing incentives for efficiency and quality (the shared savings) that reduce CMS’s financial risk. In CPC, payments in the early years were based on PBPMs. In later years, the PBPMs were reduced and the shared savings were implemented. This strategy also seems to better match payments with the different phases of transformation.

4.1.2 Learning System Features

Key Findings on Learning System Features

- In designing an initiative’s learning system, there is value in balancing the benefits of centralized oversight with the need to tailor TA on practice transformation opportunities for the circumstances of different regions or individual practices. The TA offered within an initiative needs to be tailored for differences in practices’ level of transformation.
- Practices especially value one-on-one direct assistance and peer-to-peer learning opportunities.
- It is important for TA to start early, but mid-course adjustments are often necessary.
- Health IT and use of practice feedback reports are important topics for TA.

Practices trying to adopt the PCMH model or other advanced primary care concepts face a daunting set of tasks, such as adapting their EHRs, managing organizational change, hiring new staff and integrating new functions like care coordination into the practice’s workflow, and collecting, reporting, and analyzing quality data. In addition, initiative practices also faced challenges in complying with administrative requirements (such as for NCQA PCMH recognition). To help participants overcome these challenges, these initiatives provided a range of technical assistance (TA) and other educational support to initiative sites, which can collectively be referred to as the learning system for each initiative. We identified lessons learned for three areas concerning the learning system:

- Organization and management of the learning system
- Modes for delivering TA
- TA topics

4.1.2.1 Organization and Management of the Learning System

In designing an initiative’s learning system, there is value in balancing the benefits of centralized oversight with the need to tailor TA on practice transformation opportunities for the circumstances of different regions or individual practices. One key design issue is whether an initiative’s learning system should be centrally managed or decentralized. A centrally managed learning system enables consistent national messaging across large, multi-site initiatives. One key informant noted that maintaining oversight of the learning system also gives CMS a “line of sight” into the issues practices are facing as they implement the initiative and their level of progress on transformation. A centralized learning system, however, may be less flexible in tailoring TA for the circumstances or needs of specific regions or sites. A less centralized approach that includes regional TA providers can also build and leverage stronger local relationships between practices and TA staff, as well as aligning with other regional initiatives that may be underway. Decentralized learning activities can also align with practice characteristics that cross regions, such as peer-to-peer learning for rural practices, small practices, or practices that use specific EHR vendors.

The CPC initiative took a hybrid approach to this centralization issue. CMMI staff and an overall prime contractor working for CMMI were responsible for the overall learning system, but the national prime contractor also worked with subcontractor organizations that served as regional learning faculties (RLFs) for each of the seven CPC regions. Topics that required a consistent national message, such as training on the administrative requirements of the initiative (e.g., data reporting, milestone schedules) or use of standardized data reports or other tools provided by CMS were delivered by central staff via webinars and conference calls or by regional faculty who presented standardized materials. The regional faculty provided the more in-depth, tailored learning that focused on practice transformation opportunities through a combination of peer-to-peer learning collaboratives and one-on-one practice coaching. Over time CPC also decentralized learning activities with cross-regional collaborations known as “action groups.”

TA also needs sufficient tailoring to account for different practices’ level of transformation within an initiative. For example, more advanced practices in two MAPCP states (RI and MI) reported that the technical assistance offered in their states was not useful for them and took too much time away from other practice responsibilities (Nichols et al., 2017).

A second management challenge for the learning systems in these initiatives is that it was often hard for TA recipients to access all of the learning opportunities that were available. Key informants mentioned a user-friendly website as an important tool to help practices navigate the learning system’s offerings. This website should include a well-organized library that makes it easy to find TA topics and materials. In both CPC and SIM, CMS revamped the website supporting the learning system to improve its functionality and organization.

It is important for TA to start early, but mid-course adjustments are often necessary. Another finding regarding learning system management is that it is important to start TA early in an initiative's implementation. For example, in the FQHC initiative, evaluation reports and key informants interviewed for this study both noted significant delays in organizing and delivering many of the TA activities. These delays may have accounted for the slow pace in reaching the initiative's goals during its first two years.

Beyond starting early, however, the learning system needs to stay nimble and avoid getting locked into a particular structure or curriculum plan, which is important for several reasons. First, the initiative convener and the local participating practices may differ in their views on what types of TA or topics are most important, in part because different practices may start the initiative at different stages of transformation. Second, the practices' own views on what they need from the learning system are likely to change over time, as they get further into implementation of the initiative and practice transformation. Third, those who are overseeing or delivering the TA may realize that some aspects of the learning system are not working as effectively as expected and need adjustment. Notably, although the circumstances differed, CMS ended up changing the TA contractor or significantly revising the contractor's role during implementation of both initiatives for which CMS was responsible for sponsoring major TA delivered to practices (CPC and FQHC).

4.1.2.2 Types of Technical Assistance

Another challenge is deciding how to deliver learning activities because practices value both one-on-one direct assistance and peer-to-peer learning. Delivery modes for TA include making tools and resources available online or in print, conference calls, webinars, peer-to-peer learning collaboratives, and one-on-one coaching of individual sites. In the CPC and FQHC initiatives, which included CMS or its contractors providing direct, one-on-one assistance to practices, evaluation reports and key informants interviewed for this study indicated that many practices highly valued this type of TA, especially on-site. However, this mode of TA is more resource-intensive, which may pose a challenge for scaling or sustainability. As a result, one-on-one assistance often focused on practices that needed more help, while it was less commonly provided to more advanced, higher-performing practices. The value of one-on-one assistance also depends heavily on the quality of the specific individuals delivering the TA (e.g., the practice coaches).

Peer-to-peer learning collaboratives were also popular with many participants. Key informants viewed CPC's face-to-face gatherings of each region's participating practices as useful not only for explicit learning but also for building relationships, although the time and expense of face-to-face gatherings led to a reduction in frequency from three in the CPC's second year to two in the third year. The CPC evaluator also found that most practices valued these all-day face-to-face meetings, but many practices send non-clinician staff rather than clinicians, perhaps due in part to the time involved for attendance. There were also some CPC practices that expressed more support for learning from experts on the regional learning faculty rather than their peers (Peikes, Anglin, et al., 2016). Key informants also mentioned the importance of facilitating more informal opportunities for peer-to-peer learning, such as online and telephonic discussions.

4.1.2.3 Technical Assistance Topics

Important topics for TA included health IT and the use of practice feedback reports. Practices received technical assistance on a wide variety of topics across the initiatives, including meeting administrative requirements of the initiative (e.g., data reporting by the practices), PCMH recognition requirements, leadership and organizational change, practice-level quality improvement, primary care staff development and integration of new staff, care coordination, health IT, and use of practice feedback data. Assessing the importance of specific TA topics is difficult because practices differed widely in their circumstances. As a result, key informants in this study repeatedly mentioned the importance of tailoring TA topics for different recipients' levels of expertise. Informants also noted that practices' needs for TA topics can change over time as they make progress on transformation.

Notwithstanding these issues, key informants identified health IT as an important TA topic, which is consistent with findings in the evaluation reports that many initiative sites faced challenges in using their EHRs and in obtaining near real-time knowledge of patients discharging from an inpatient stay or an emergency department visit. Key informants also recommended more learning activities related to use of practice feedback reports that were provided in the initiatives; however, informants had mixed views on whether more TA would actually increase use of these data reports (see section 4.3).

Key informants also emphasized that TA recipients want tangible tools and practical, concrete options or solutions for issues, rather than TA presentations that are more conceptual in nature. As one key informant noted, TA needs to be grounded in an operational context to be worth the recipients' time.

4.1.3 Provision of Data to Practices

Key Findings on Provision of Feedback Data

While many practices perceived some benefit from the feedback reports provided by CMS, key challenges that limited the value of the data include the following:

- Practices often viewed quarterly, claims-based data as not timely enough to be actionable for care management.
- Some primary care physicians perceived their feedback reports as not actionable because results could be driven by factors beyond their influence (e.g., the behavior of specialists and hospitals).
- Privacy barriers to sharing mental health and substance abuse data posed a content problem.
- Practices want aggregated, multi-payer data.
- Review of external data reports is not part of a physician's normal workflow.
- Practices did not receive enough training in how to interpret and use the data provided.

Advanced primary care practices require data to inform their efforts to deliver care more effectively. In addition to practice-level data on quality, utilization, and costs, patient-level data may help practices in coordinating care, especially for high-risk patients, as well as supporting broader population health efforts. While our review of the six initiatives indicates there is value

in providing such feedback data to participants, there were also multiple challenges that limited the value of the data provided.

4.1.3.1 Use of Data in Practice Transformation

Conveners anticipated that providing data to practices would be an important tool for primary care transformation. Participating practices in the six initiatives typically received a range of data to provide feedback on their performance in the initiative and to support transformation activities. Many practices found some value in these reports, but usage varied greatly across initiatives. For example, a 2015 survey of CPC practices found that 77 percent of practices reviewed quarterly practice-level Medicare feedback reports most or all of the time and 53 percent reviewed Medicare patient-level reports most or all of the time (Peikes et al., 2016). Midway through the CPC initiative, CMS gave the feedback data a more explicit role in transformation activities by establishing a milestone requirement that practices use at least one payer's feedback report to identify an issue to target for cost or utilization savings, along with an increased emphasis on the use of feedback reports in the learning system's delivery of technical assistance to CPC practices. Among CPC practices that had reviewed the Medicare feedback reports as of 2015, 90 percent said the practice-level reports were somewhat or very useful and 82 percent said the patient-level reports were somewhat or very useful (Peikes, Anglin, et al., 2016).

Conversely, in the FQHC demonstration, sites began receiving quarterly feedback data reports halfway through the initiative (quarter 6), but less than 20 percent of sites accessed these reports in the first two quarters they were made available. After site visits by the evaluation contractor made clear that most sites were initially unaware of the feedback reports or did not perceive them as valuable, CMS and the learning system contractors increased their efforts to promote the feedback reports, and by the last quarter of the initiative 86 percent of the sites had viewed at least one report. FQHCs generally found a feedback report addressing readiness for PCMH recognition to be helpful, but most sites did not find the other feedback reports, focused on costs and utilization, to be particularly useful, in part because the Medicare data only reflected a small share of the FQHC patients (8 percent on average).

The experience in MAPCP varied substantially across the states. Among the five states for which CMS provided feedback reports, one state (NY) had more than 90 percent of practices accessing the reports, but in the other four states (ME, PA, RI, and VT) less than 35 percent of practices downloaded their reports in any given quarter through the end of 2014 (Nichols et al., 2016). However, at least two of these states with low usage of CMS reports (RI and VT) were providing their own reports to MAPCP practices.

Of 31 initiative settings included in the ESF questionnaire that we asked evaluators to complete (covering FQHC, IAH, seven CPC regions, eight MAPCP states, six SIM states, and eight of the HCIA PCR awardees), 26 settings (84 percent) provided practices with feedback data reports—typically on a quarterly basis. The evaluators for 16 of these 26 settings (62 percent) judged that practices adapted or changed their transformation activities based on these reports to a moderate extent (in 14 settings) or to a great extent (in two settings—MAPCP VT and HCIA TransforMED).

4.1.3.2 Challenges in Providing Data to Aid Practice Transformation

Notwithstanding these indications that many practices perceived some benefit from their feedback reports, the initiative evaluation reports and our key informant interviews highlighted a range of challenges that limited the value of the data:

- **Practices often viewed quarterly claims-based data as not timely enough to be actionable for care management.** Data were typically produced on a quarterly basis, with a time lag of three to six months to account for claims processing and data production. While some practices found the claims-based data useful to inform practice-level changes (as noted above), key informants we interviewed indicated that practices want data on a near real-time basis to support care management or other clinical activities. The CPC and MAPCP evaluation reports also noted practices' concerns about time lags in their feedback reports (Peikes, Anglin, et al., 2016; Nichols et al., 2016). Practices particularly want near real-time knowledge of patients discharged from an inpatient stay or an emergency room visit, which is not a goal that can be met with claims-based reports. Several states in SIM and MAPCP addressed this issue by implementing local or statewide health information exchanges that provide such discharge notifications to primary care practices, and many other initiative participants eventually addressed this issue by making their own notification arrangements with local hospitals (Gavin et al., 2016; Nichols et al., 2016). Time lags associated with claims-based data also made it difficult for practices to reconcile data from CMS and other payers with more current data generated from their own EHRs. For example, the MAPCP evaluation found practices were frustrated with inconsistencies between patient risk scores assigned in quarterly claims data reports versus the practices' own view of patients needing care coordination based on more current EHR data and the clinicians' knowledge of their patients (Nichols et al., 2016).
- **Some primary care physicians perceived their feedback reports as not actionable because results could be driven by factors beyond their influence,** such as the behavior of specialists and hospitals (Peikes, Anglin, et al., 2016). In an interesting attempt to focus on content that primary care physicians would view as actionable, the CPC initiative is implementing a report detailing patients' usage of specialty providers. As a key informant noted, this question is something claims data are well suited to answer, and the information does not need to be as timely to be actionable for improved care coordination and efficiency. CPC primary care physicians may learn of specialists who see a high volume of their patients, and with that information the CPC practices can contact those specialists to initiate better care coordination procedures.
- **Privacy barriers to sharing mental health and substance abuse (MH/SA) data posed a content problem for feedback reports.** A key informant involved in MAPCP noted that MH/SA conditions were often excluded from patient-level data provided to practices, but patients with MH/SA conditions are often most in need of precisely the types of advanced primary care services these initiatives are intended to support, such as care coordination. The CMS contractor evaluating the SIM Round 1 model test states described this issue as "a significant challenge" (Gavin et al., 2016,

p. 37). The SIM states have attempted to address this challenge by educating stakeholders concerning relevant privacy requirements and appropriate sharing of MH/SA data.

- **Practices want aggregated, multi-payer data.** As a key informant noted, when different payers provide their own separate data reports, practices are frustrated that they have to reconcile different formats, definitions, or results. Practices' desire for aggregated, consistent multi-payer data is particularly evident in CPC, as multiple key informants mentioned this issue in interviews for this study and the CPC evaluation contractor documents the status of data alignment or aggregation efforts in five of the seven CPC regions (Peikes, Anglin, et al., 2016). Several of the CPC regions are taking steps to address this challenge by creating multi-payer claims databases, but doing so has proven time-consuming and raises its own challenges, especially concerning inclusion of Medicare data. Building a multi-payer claims database and reporting capability typically requires the payer partners to agree on a prime contractor for the work involved, but CMS does not have the same contracting flexibility as the commercial payers.
- **Review of external data reports is not part of a physician's normal workflow.** Physicians in these initiatives face many competing priorities for their time beyond direct patient care, leaving little time or inclination to review an external data report from CMS or other payers—especially if the physician does not view the data as actionable. Some practices address this challenge by having a staff member other than a physician responsible for reviewing the data reports, but smaller practices are particularly unlikely to have a staff member with the available time and skill set for such analysis, and multiple key informants noted that larger practices are more likely than smaller practices to review the feedback data provided by CMS.
- **Key informants reported that practices did not receive enough training in how to interpret and use the data provided.** Informants had mixed views, however, on whether more training would lead to a significant increase in use of the data reports, given all the other challenges described herein. Informants also noted that the individuals involved in coaching practices vary in their capacity to teach practices how to use the data reports. This challenge potentially could be addressed with more centralized training of practice coaches concerning the feedback reports.

4.1.4 Easing Administrative Burdens Could Help Practice Transformation

The initiatives required practices to conduct a number of activities related specifically to the initiative. For example, FQHC practices had to meet the NCQA PCMH recognition standards, which require documentation of practice activities. On top of the efforts involved in actual practice transformation, the administrative requirements associated with the initiatives were often a burden. The initiatives tried to provide support for these activities (e.g., financial support for NCQA fees and TA for how to comply with initiative requirements). However, there were still many challenges.

One type of challenge was the level of documentation and reporting associated with initiative requirements. In the FQHC initiative, which required FQHCs to meet NCQA Level-3

recognition as a PCMH, FQHCs reported that the level of documentation required by NCQA was burdensome because the FQHCs had to create and document new, formalized policies to support NCQA's requirements. Many FQHCs considered these documentation efforts to be a distraction from actual practice transformation. In fact, meeting some of the NCQA recognition deadlines caused FQHCs to stop their implementation efforts as they shifted to completing paperwork. In CPC, some practices that were part of larger systems reported that CPC's unique requirements forced them to change their workflow and documentation/reporting processes, which they thought was inefficient. In the IAH initiative, practices had to submit documentation of patient care preferences and medication reconciliation in the home, neither of which were available in claims data. This caused IAH practices to have to develop methods to extract these data. CPC practices also identified the initiative's documentation requirements as a challenge.

A second challenge stemmed from the approaches used to push practices to reach the initiative's transformation goals. The CPC initiative established a list of nine milestones that practices had to meet each year. CPC practices indicated that meeting the many milestones in the initiative could be overwhelming, and they would have preferred to focus on fewer milestones (Peikes et al., 2016). Although these requirements caused some practices to leave the initiative, most CPC practices met the milestones. Establishing the appropriate set of guidelines/requirements for initiative practices is not straightforward. For example, in the early portion of the FQHC initiative, CMS did not establish the type of milestones used in the CPC initiative. This led to little early progress by FQHCs and may have been responsible for so many FQHCs not achieving PCMH recognition until the very end of the demonstration. Because much of the work to achieve NCQA recognition in the FQHC initiative was not undertaken until the third year of the initiative, practices thought that the changes then had to be made at a rushed pace.

A third type of administrative burden noted by participants in several MAPCP states was that they reported spending too much time determining which patients were eligible for enhanced services such as care management or home services. In some cases, different payers used different algorithms for identifying high-risk candidates for care management; in other cases, payers' lists of high-risk patients that were based on claims data did not match the clinical information that practices had internally from their EHRs (Nichols et al., 2016). The CPC initiative addressed this challenge by having all patients in CPC practices receive the initiative's enhanced services.

4.2 Practice-Level Facilitators

We identified some important factors that were internal to practices that helped them become advanced primary care practices: practice readiness (discussed in 4.2.1), having health IT capabilities (discussed in 4.2.2), and the ability to integrate new staff into the practice's workflow (section 4.2.3).

Practices transformed considerably during the initiatives. Before discussing the internal factors related to practice transformation, it is useful to examine the actual amount of transformation that occurred for the initiative practices. We found that in each initiative there was evidence that practices did modify their procedures and that they moved toward becoming advanced primary care practices. For example, although less than 10 percent of the FQHCs

included in the FQHC initiative had any type of PCMH recognition prior to the demonstration, 70 percent of the initiative FQHCs achieved NCQA Level-3 PCMH recognition by the end of the three-year demonstration period.

We conducted an analysis for this meta-evaluation, which analyzed the level of practice transformation during each initiative. We based this analysis on two surveys. First, we used evaluator responses to the ESF survey that we conducted as part of this study to measure changes in six PMCH domains. We measured practice transformation levels based on a simplified version of the Safety Net Medical Home Score (SNMHS) (Nocon et. al., 2012; Birnberg et al., 2011). The SNMHS defines the extent of patient-centered medical home (PCMH) adoption with respect to the following six domains:

- Access and communication
- Care management
- External coordination
- Patient tracking and registry
- Test and referral tracking
- Quality Improvement

The evaluators responding to the ESF survey assessed each domain before the initiative and then in 2015, at the setting level (i.e., for each CPC region, MAPCP state, HCIA awardee, and FQHC overall). To minimize response burden for evaluators, the ESF questionnaire described the content of each domain and asked respondents to rate the extent to which each setting had achieved the domain goals using a scale ranging from 1 = not at all to 4 = to a great extent. Respondents were first asked to make assessments of each domain before the demonstration began and then again after the demonstration had ended. All domain ratings were converted to 0 (lowest possible adoption level) to 100 (highest possible level) scores, and the domain scores were averaged to derive the total SNMHS score. We were able to calculate scores for 16 settings (seven CPC regions, FQHC, two HCIA awards, and six MAPCP states). We found that the practice transformation score increased at 14 of the 16 settings and that the scores increased by an average of almost two-thirds for the practices in these 16 settings. We found the greatest percentage increase in scores for the settings with the lowest levels of pre-demonstration scores. As a second method, we used assessments done by the CPC evaluators using the PCMH-Assessment (PCMH-A). The CPC evaluators measured pre-demonstration and post-demonstration scores on the PCMH Assessment and found that practices in all seven CPC regions increased their score substantially (a range of 36–76 percent, with a median increase of 47 percent.) The findings of the two analyses indicate that, using the measures in the SNMHS and PCMH-A surveys, practices in the initiatives changed considerably during these initiatives and made progress to becoming advanced primary care practices.

4.2.1 Readiness

Key Findings on Readiness

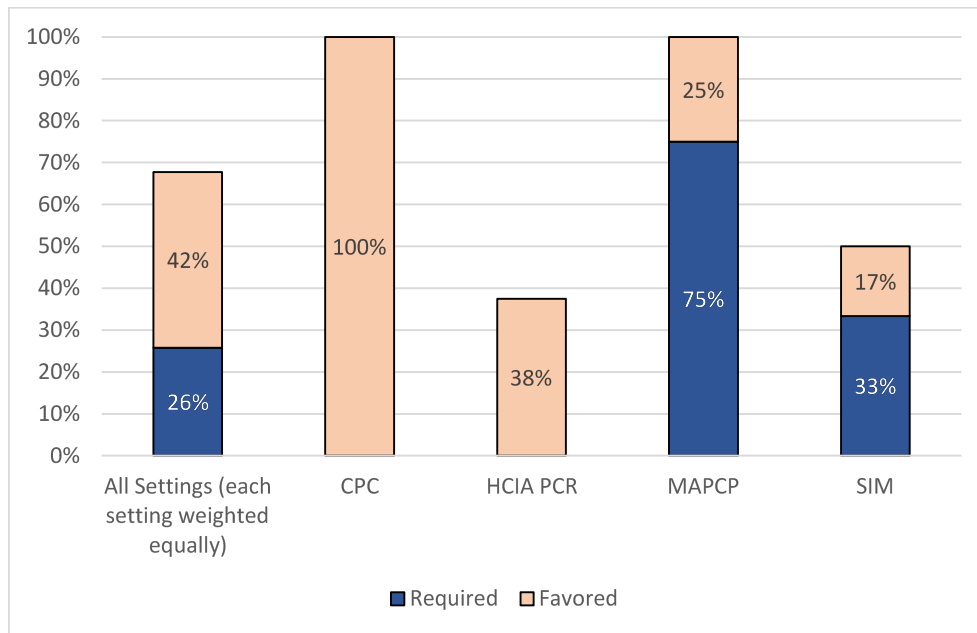
- Experience implementing practice transformation activities facilitated initiative implementation.
- Effective and committed practice-level leadership helps build a collaborative work environment and facilitates implementation.
- Smaller practices will likely need additional transformation support and may need alternative milestones as they may lack the staff, resources, and capacity to manage implementing transformation.

Practices’ experience with health care transformation often gave them an advantage in implementing the various initiatives. Practices with practice transformation experience knew what obstacles to expect and needed less time to become operational. For example, many HCIA awardees were able to expand on existing pilot programs and already had staff and infrastructure in place. CPC evaluators reported that prior to the initiative several practices formed dedicated quality improvement teams, incorporated shared decision-making tools, or planned to integrate care coordinators. Key informants characterized these practices as having “an implementation climate conducive to new work,” and this culture facilitated early implementation success. This success was a result of staff experience with quality improvement efforts, new technology, and new staff roles. In addition, practices that had relationships with community organizations found that this also facilitated care coordination.

A lack of experience hindered primary care transformation. Practices with little or no experience in primary care transformation did not fully understand what primary care redesign entailed, took longer to become operational, needed additional technical assistance, and set overly ambitious goals. In addition, these practices were often unprepared for the requirements, terms, and conditions of a CMS initiative. One HCIA key informant said he wished they had vetted awardees for inclusion in the initiative based on experience or capacity because that experience could have helped practices avoid pitfalls, set realistic goals, and collaborate with partners to implement their plans more efficiently and effectively. On the other hand, less experienced practices may be less invested in specific intervention components and thus be more willing to change.

Most practices selected for the initiatives had already undergone some primary care transformation. In selecting treatment practices for the initiatives, we found that four conveners and nearly 70 percent of settings required or favored for inclusion those practices that had experience with transformation, as illustrated in *Exhibit 4-2*. For example, nearly all practices in the MAPCP initiative had some level of PCMH recognition prior to selection. The FQHC initiative did not explicitly seek out practices that had experience, but many practices were already working toward NCQA recognition and improving access to care (e.g., same-day services) when they joined the initiative. These practices were closer to achieving PCMH status compared to those practices without prior work toward transformation. This means that initiative practices should have been generally ready for transformation to advanced primary care practices.

Exhibit 4-2. Percentage of Settings that Required or Favored Practices with NCQA/State Accreditation



Source: Questionnaire completed by the CMS evaluation contractors for this study

Note: Neither IAH nor FQHC required or favored practices with transformation experience for inclusion.

Setting refers to state for SIM and MAPCP, region for CPC, and awardee for HCIA. Total settings by initiative: CPC (7), FQHC (1), HCIA-PCR (8), IAH (1), MAPCP (8), and SIM (6)

Practice “champions” improve staff buy-in and facilitate transformation. A successful implementation is one where the planned services were consistently delivered to the intended target population at the intended level of quality and intensity (Glasgow, Vogt, & Boles, 1999). Many key informants reported a practice champion was key to successful practice transformation. In-depth interviews conducted for the CPC evaluation identified five characteristics of successful physician champions (Peikes, Ghosh, et al., 2016):

- Enthusiastic about the initiative
- Helps guide change
- Listens to the suggestions of the primary care team on workflows
- Acknowledges progress made
- Gives feedback to the team

A committed practice champion can also encourage staff buy-in. An IAH key informant reported that a practice champion’s buy-in has a trickle-down effect and the “*more engaged the primary lead was, the more engaged the staff were.*” The HCIA evaluation also found that effective leaders empower their staff and encourage team communication. Physician champions

are needed to lead their practices' transformation efforts and to foster a more collaborative and communicative work environment that builds staff buy-in and drives the initiative forward.

The FQHC evaluation found that when the practice champion was also an organization or practice leader, he or she had the additional capacity to allocate supplemental resources (time, space, and materials) for transformation activities. For FQHCs, this strong and consistent leadership support was foundational to achieving NCQA Level-3 PCMH status by the end of the demonstration (Kahn et al., 2016).

Smaller practices faced staff, resource, and capacity challenges implementing transformation. Key informants from two initiatives reported that transformation was difficult for smaller practices because they had fewer staff and limited financial resources. Although one IAH key informant reported smaller practices can change their practice and track patients more easily than large practices, they cannot hire staff quickly or adjust clinicians' hours as easily as larger practices. As discussed in section 4.1.4, fulfilling initiative requirements (e.g., documentation) may also require capacity that smaller practices do not have. Key informants noted that smaller practices are also less likely to have the capability to understand and analyze feedback data (as discussed in 4.1.3). A CPC key informant said that for smaller practices *"just running the business takes effort and you don't have the economies of scale [in transformation] that bigger practices do."* Small practices sometimes joined larger health care systems that had the staff and budget to invest in transformation. However, joining a health system led to a loss of autonomy when the health system did not successfully engage them. One CPC key informant suggested that mid-sized practices may be successful because they have sufficient funding and staff to invest in transformation, but they are able to retain their decision-making power.

4.2.2 Health IT

Key Findings on Health IT

- Practices with high-functioning health IT have tools that support a smoother transformation.
- Health IT solutions need to evolve to meet the demands of practice transformation, including improved care coordination, care management, and payment reform.
- Implementation of EHRs is a frequent barrier to practice transformation.

High-functioning health information technology (health IT) facilitates practice transformation. Health IT refers primarily to EHRs, personal health records, and health information exchange networks and functionalities (The National Alliance for Health Information Technology, 2008). High-functioning health IT, a requirement of NCQA certification, is intended to give practices the necessary infrastructure for many practice transformation tasks, such as documentation, referrals, scheduling, population management, and care coordination. However, evaluators identified challenges related to implementing high-functioning health IT, including an extended planning, testing, and rollout process.

One-third to one-half of FQHC practices reported "EHR system functionality" as a facilitator to practice transformation (Kahn et al., 2016, p. 97). EHRs are a key health IT component, and both FQHC and CPC informants commented that practices would be more able to transform if they possessed basic EHR capacity. For FQHC practices, having pre-

demonstration EHR system functionality predicted success in earning NCQA Level-3 PCMH recognition by the end of the demonstration (Kahn et al., 2016). CPC practices varied in their use of EHRs, and the CPC evaluation suggested that the practices that had already satisfied federal EHR meaningful use requirements had fewer changes to make when they joined the CPC initiative. As one key FQHC informant noted: *“A majority of the [FQHC practices] mentioned that having an EHR system that had been in place and fully functioning for two years prior to PCMH changes assisted in the uptake of providers to the new processes of care coordination, referrals, population management and scheduling. Having the EHR in place laid a foundation for providers to have a stable interface and to be comfortable with the variety of screens and how to manipulate the EHR for PCMH documentation.”*

One key informant suggested working with EHR vendors directly to better support practices in transformation. Interoperability of EHRs, particularly for smaller practices, was a challenge. According to one evaluator, “The lack of interoperability of EHRs across practices ties the hands of managing care.” CPC created “affinity groups” connecting practices with vendors to discuss issues and get support. A CPC key informant thought this idea could be pushed further by organizing practices around their health IT vendors “because ... as a group, practices might be able to push a vendor to do more to help them than [as] individual [practices].” The CPC+ expansion of CPC includes a provision for EHR vendor support of participating practices.

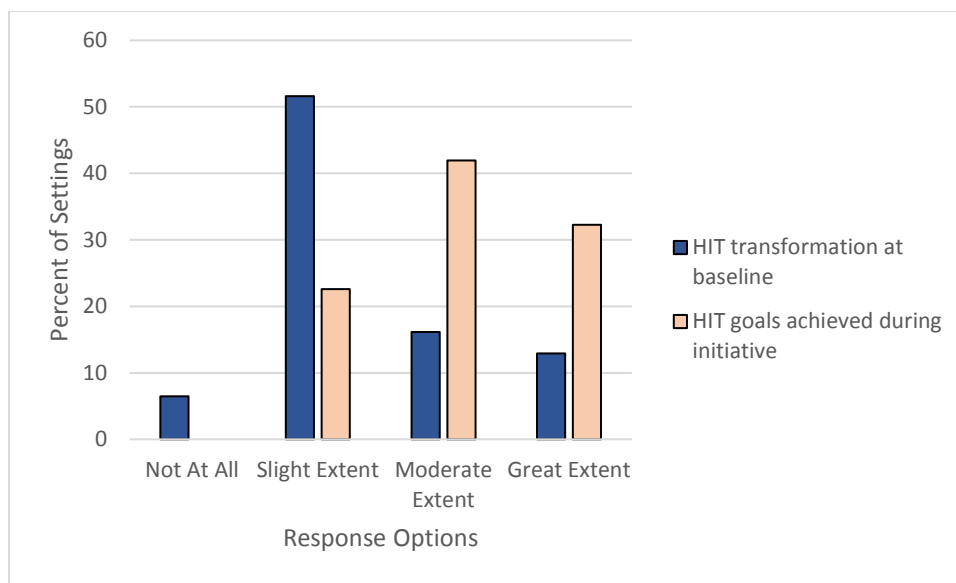
Health IT infrastructure may be an indicator of other features that support transformation and practice readiness for change. The CPC evaluation noted that having high-functioning health IT often signals the presence of other transformation supports, such as committed leaders and an organizational climate open to change. However, as discussed below, even practices with reasonable health IT capacity at baseline often struggled to implement practice transformation activities.

Health IT needs to evolve to meet primary care transformation goals. We found that, prior to the start of the six initiatives, most initiative practices were engaged in health IT transformation to only a “slight extent” (*Exhibit 4-3*). During the initiative, practices invested more time and more financial resources than expected to adapt health IT systems to address transformation activities. Although some practices were more advanced with respect to EHRs, health information exchanges, or in-house data management expertise, key informants across all the initiatives reported current health IT systems were not designed to support primary care transformation activities such as care management, care coordination, and payment model reform. Some of the specific problems include the following:

- **There were challenges using EHRs for care management.** EHRs are designed for visit-based, in-person care by an independent licensed provider, while care management often involves pre-planning or between-visit patient interactions. Key informants reported that practices would either invest unanticipated time and money in creating templates for care management or use two different electronic systems for clinical care and care management. Some practices also struggled with implementing risk stratification algorithms in EHR systems to support care management for the most complex patients (Taylor et al., 2015).

- Current health IT infrastructure does not support *care coordination*, because of a lack of interoperability among EHRs or lack of a functioning health information exchange. Providing a care summary that includes all of a patient’s specialty providers proved difficult when different providers used different EHR vendors that could not share information. Likewise, many practices could not perform timely hospital discharge follow-up, because they did not have access to the hospital discharge information. Practices usually employed cumbersome workarounds, such as faxing paper records and manually entering information into the EHR.
- **Inadequate health IT also made it difficult to implement an *alternative payment model at the practice or state level*.** Many practices in CPC had EHRs that could not aggregate results to the practice level (Taylor et al., 2015). All MAPCP states struggled with poor data quality in statewide clinical registries or in EHRs, which complicated patient attribution and payment (McCall et al., 2015b). Finally, the CPC evaluator reported that integrating shared decision making tools into EHRs and clinical workflow was challenging (Taylor et al., 2015).

Exhibit 4-3. Extent to Which Settings Achieved health IT Transformation



Source: Questionnaire completed by the CMS evaluation contractors for this study

Across the initiatives, we found that these EHR limitations had a large impact on *vulnerable patient populations*, such as those who receive long-term care, behavioral health, or substance abuse services. These patients often have the greatest need for care management and care coordination. Many providers who care for these patients do not use EHRs (e.g., home health aides, mental health providers, nursing home staff), and EHRs are often not designed to care for these populations (Gavin et al., 2014). Additionally, federal restrictions prohibit or limit sharing data from substance abuse or mental health providers (Institute of Medicine, 2006), preventing the primary care team from having a comprehensive understanding of a patient’s health.

Despite challenges and health IT shortcomings, almost 75 percent of settings achieved high-functioning health IT transformation goals to a moderate or great extent. At baseline, evaluators estimated most settings (e.g., states, CPC regions, HCIA enrollees) had achieved only slight health IT transformation (the blue bars in *Exhibit 4-3*), though when queried on the extent to which each setting subsequently achieved health IT goals (the orange bars), many evaluators reported “to a moderate extent.” Practices in the early stages of EHR adoption transitioned from paper to electronic charting or adapted the EHR to meet practice needs. For example, one IAH practice adapted its EHR to push home visit reminders. Practices with more developed systems invested in technology to perform advanced primary care functions, such as population management systems that integrate with the EHR, or created risk stratification software that allowed care managers to identify priority patients. Several settings also developed health information exchanges promoting improved care coordination, such as DocSite in MAPCP Vermont or the North Carolina Case Management Information Systems in MAPCP NC (McCall et al., 2015a).

4.2.3 Integrating New Staff

Key Findings on Integrating New Staff

- Hiring new staff is a common approach to support practice transformation.
- Allowing sufficient time for staff to perform practice transformation activities and preparing for staff turnover facilitates practice transformation.
- Clearly defined roles for newly hired staff may lead to improved hiring practices, retention, and facilitate practice transformation.

Most transforming practices hired new staff. Hiring new practice staff or partnering with organizations outside the practice promoted practice transformation across initiatives. Frontline evaluators for CPC, HCIA, and MAPCP estimated that more than 50 percent of practices within each setting hired new front-line staff to support the initiative; fewer practices in IAH, FQHC, and SIM engaged in new hiring. Practices often hired clinical staff to expand hours or handle higher volumes and hired non-clinical staff to incorporate innovations like care management or care coordination into clinical workflow. For example, some practices hired dedicated staff to register patients for an online patient portal or to respond to the increased volume of patient phone or email communication (Peikes, Ghosh, et al., 2016). The CPC evaluation found that 90 percent of CPC PBPM payments were used for labor costs associated with physicians, RNs, and care coordinators (Peikes, Anglin, et al., 2016).

Some initiatives provided practices with external staff on a shared or embedded basis, particularly for several of the MAPCP initiative states and HCIA awardees. Key informants for MAPCP described the critical role that community health teams played in managing care outside of office visits, especially for smaller practices where hiring a care coordinator was not feasible. Community health teams were able to organize care and track patients across providers, identify ED frequent flyers, and alert practices when patients were hospitalized three times. Community health teams were also helpful in assisting practices in understanding and using feedback reports. On the other hand, community health teams did not have access to EHR records, and some had difficulty accessing practice-level data. Moreover, workarounds were often required to pay

community health team workers because community health teams are not considered providers or employees of the practices themselves that Medicare is able to bill.

Staff turnover impeded implementation progress. Turnover or unfulfilled roles were a significant problem for more than 50 percent of practices in FQHC, IAH, MAPCP, SIM, several HCIA awardees, and some CPC states (CO, OK, OR). For example, according to key informants, provider and staff turnover in FQHCs may have negatively affected the development of cohesive care teams and the stable empanelment of patients to providers (Kahn et al., 2016). Practices in CPC and FQHC that experienced staff turnover had trouble achieving milestones, fulfilling reporting requirements, or using feedback reports because procedural and content knowledge were lost with every departure.

Practices struggled to integrate new staff members to support primary care innovations. As one key informant reported, practices intending to provide care management with unlicensed staff frequently also had to hire skilled staff such as nurses or social workers to work with medically and socially complex patients. In SIM, key informants described how multiple providers and payers were incentivized to coordinate care through several different avenues. As a result, some practices ended up working with multiple care coordinators and were faced with “*coordinating the coordinators*,” highlighting the need to ensure transformation efforts are complementary and not duplicative. Because health care delivery innovations often add staff with roles distinct from existing staff members, conscious integration of new staff into the clinical workflow aids transformation. Defining clearly the activities of new staff members and refining those roles based on the goals of the initiatives may lead to improved hiring practices and retention, promoting transformation of the clinical workforce to better meet patient needs.

4.2.4 Identifying a Practice’s Patient Population

Practices in the initiatives wanted to know which patients they were responsible for and for which patients they would receive financial support.

The number of patients attributed to practices was far less than practices expected. This was concerning to practices because attribution links beneficiaries to the particular practices so that the practices can receive payments. Practices consistently identified treating more Medicare patients in their practice than ended up being attributed to their practice as a concern (McCall et al., 2015a). This occurs because patients can only be attributed to one practice, but some Medicare patients see multiple PCPs at multiple practices. Newer patients will often be attributed to their previous PCP’s practice. For Medicaid patients, using claims files to link patients to practices is even more challenging because of their intermittent eligibility. MAPCP and SIM evaluation reports also highlighted a concern about patients who live in a demonstration state only part of the year, causing inconsistent attribution to practices (Gavin et al., 2014; McCall et al., 2015a). In net, a MAPCP key informant reported that often as much as 40 percent of patients that a practice expected to be attributed were not. As a result, practices received often lower than expected financial support to invest in practice transformation (Taylor et al., 2015).

Many practices found this uncertainty in identifying and attributing eligible patients to specific practices to be challenging (Nichols et al., 2017). Practices unsure of patients attributed to their practice are also unsure of the financial support to expect. Providing more education to practices on attribution policies may create more transparency and provide more certainty to practices.

5. How do the Initiative Impacts Compare and Contrast?

To compare and contrast the evaluation findings, we used meta-analytic and descriptive methods to compare the effects of the different primary care initiatives and settings on the four core outcomes identified by CMMI: medical expenditures without fees, outpatient ED visits, hospital admissions, and 30-day readmissions. We reviewed the evaluator reports to collect the latest or final reported difference-in-differences (DiD) estimates for a setting as of June 2017. The DiD estimate represents the average regression-derived difference in performance achieved by each setting relative to each setting's comparison group.

We examined the results for the 22 settings for which we had evaluator-reported data on cumulative DiD estimates: eight MAPCP state settings (NY, RI, VT, NC, MN, ME, MI, PA), seven CPC regional settings (AR, CO, NJ, NY, OH/KY, OK, OR), six HCIA-PCR awardee settings (CareFirst, Denver Health, Finger Lakes Health Systems Agency, PeaceHealth, Sanford Health, TransforMED), and one overall setting for FQHC. The cumulative DiD estimates for all settings except HCIA-CareFirst (for whom we had only the first two intervention years) are based on at least three intervention years. We omitted IAH and the six SIM state settings (AR, ME, MA, MN, OR, VT) from these analyses due to insufficient data. We also omitted two HCIA awardee settings from our quantitative analyses and the QCA analysis; while the evaluator for Pacific Business Group on Health (PBGH) and Cheyenne/WIPH estimated the program impacts, the evaluator cautioned against drawing conclusions about the impacts due to selection problems (PBGH) and failure to implement the innovation as planned (WIPH).

This report does not present initiative-level results because the evaluation contractors reported DiD results using different metrics (e.g., per member per month) that had to be standardized to a common metric or, in the cases of MAPCP and HCIA, the evaluators did not report initiative-level findings.

We first examined two key questions that were of substantive interest and were also important to our methodology for comparing initiative results. First, does performance on the four core outcomes differ significantly by initiative? The answer to this question determined whether we needed to account for initiative in our meta-regression models. Second, does performance vary over time? The answer to this question determined whether a focus on the cumulative DiD effect size is warranted, rather than year-specific results. As presented in further detail below, we found that for each of the core outcomes the estimated results at the initiative level did not differ significantly from each other among the four initiatives. We also found that the year-by-year results are strongly correlated. These results support our focus on setting-level, cumulative performance estimates in our subsequent analyses. We also assess the overall impact of these initiatives on each of the four core outcomes (as a grand mean of all 22 settings in our analysis) using formal random-effects meta-analysis, which assumes that settings differ in substantive ways and that a common effect cannot be assumed. We close the chapter by addressing whether, and to what extent, performance varied by setting. More information on the DiD estimates used and meta-analytic methods can be found in *Appendix D*.

5.1 Differences in Outcomes by Initiative and Over Time

At the initiative level, performance on the four core outcomes rarely differed significantly by initiative. We used random-effects meta-regression to compare cumulative

three-year outcomes by initiatives (*Exhibit 5-1*). We found that the cumulative three-year impacts on each of the core outcomes, at the initiative level, in almost all cases were not significantly different for FQHC-APCP, HCIA-PCR, and MAPCP relative to CPC as the reference category. The sole exclusion at the initiative level was that MAPCP performed significantly better in 30-day readmissions than CPC relative to comparators (11.9 fewer 30-day readmissions than CPC, $p=0.06$). This sole significant difference in performance at the initiative level allows us to ignore initiative-level features and supports our focus on setting-level features that influenced performance on the four core outcomes. How and which features affected performance are described in Chapter 7.

Exhibit 5-1. Meta-regression Results Comparing the Four Core Outcomes with CPC, by Initiative (n = 22)

Initiatives	Cumulative Year 3 Difference-in-Differences Outcomes							
	Expenditures without fees		Outpatient ED visits		Hospital Admissions		30-Day Readmissions	
	Coef.	90% CI	Coef.	90% CI	Coef.	90% CI	Coef.	90% CI
CPC	ref.		ref.		ref.		ref.	
FQHC-APCP	253.0	(-140.7 to 646.6)	34.2	(-13.4 to 81.8)	12.1	(-3.2 to 27.3)	-1.1	(-11.0 to 8.8)
HCIA-PCR	150.5	(-107.5 to 408.5)	10.1	(-17.0 to 37.2)	6.5	(-3.8 to 16.8)	2.8	(-3.0 to 8.5)
MAPCP	130.2	(-125.5 to 386.0)	14.4	(-12.2 to 41.1)	11.9	(1.8 to 22.0)	2.6	(-5.2 to 10.4)

Notes: CI = confidence interval; Coef = coefficient; ED = emergency department; ref.= reference

Models included 22 settings (7 from CPC, 1 from FQHC-APCP, 6 from HCIA, and 8 from MAPCP).

Intervention effects did not tend to improve over three years. Over time (Years 1–3), setting performance was strongly correlated from one year to the next, although the correspondence of estimates of Year 1 with Year 3 performance declines somewhat relative to adjacent years (Year 1 with Year 2 or Year 2 with Year 3; see *Exhibit 5-2*). In general, settings that were high performers remained high performers, while those that were performing at a lower level in Year 1 did not necessarily improve greatly by Year 3. Because of these similarities between annual performance effects, we focused our analyses for the core outcomes, below and in Chapter 7, on the cumulative three-year impacts. Additional details on year-by-year results on each of the core outcomes are presented in *Appendix D (Exhibits D9 through D12)*.

Exhibit 5-2. Correlations of Setting Performance for Year 1, Year 2, and Year 3 Outcomes

	Expenditures without fees		Outpatient ED Visits		Hospital Admissions		30-Day Readmissions	
	Year 1	Year 2	Year 1	Year 2	Year 1	Year 2	Year 1	Year 2
Year 2	0.584		0.765		0.584		0.517	
Year 3	0.435	0.680	0.766	0.793	0.310	0.791	0.331	0.457

Note: Correlations are weighted by the inverse of the sampling variance (average of Year 1, Year 2, and Year 3). All correlations are significant at $p < 0.10$ except the correlation between Year 1 and Year 3 for hospital admissions ($p = 0.17$) and between Year 1 and Year 3 for 30-day readmissions ($p = 0.14$). Twenty-two settings had estimates for Year 1 and Year 2; 21 settings had estimates for Year 3 (HCIA-CareFirst did not report Year 3 DiD impact estimates).

5.2 Differences in Outcomes by Setting

Performance on the four core outcomes varied significantly by setting. None of the results for the four core outcomes were statistically different from zero ($p < 0.10$) when considering all 22 settings collectively (the grand means for settings within FQHC, CPC, MAPCP, and HCIA-PCR, in *Exhibit 5-3*). However, given the findings above that impacts among the initiatives generally did not differ significantly at the initiative level or over time, we instead focused on performance on the core outcomes at the setting level. For each of the core outcomes, we see significant variation (heterogeneity) in performance among the settings, both within and across initiatives. *Exhibit 5-3* presents a summary of the cumulative three-year effects (see also *Appendix D, Exhibits D-13* through *D-16*). Using a two-sided test, significant changes in expenditures or utilization are shown in varying shades of green (reductions) and red (increases). The corresponding darker shades indicate significant changes in expenditures or utilization at the $p < 0.20$, < 0.10 , and < 0.05 levels, respectively. Null findings are white.

Of the 22 settings, more than two-fifths experienced improvement on one or more of the four core outcomes, while half returned a worse outcome relative to their comparator at a 10 percent significance level. For results on the four core outcomes among the 22 settings, at $p < 0.10$:

- Four settings reduced expenditures, while four increased expenditures.
- Four settings reduced outpatient ED visits, while five settings increased them.
- Hospital admissions fell in four settings and increased in three.
- Thirty-day readmissions fell in two settings and increased in two.

Overall, across the 22 settings and four outcomes summarized, relative to comparison groups at the $p < 0.10$ level, 68 percent of the results showed no effect, 16 percent showed improvements in outcomes (11 percent at $p < 0.05$), and 16 percent experienced worse outcomes (10 percent at $p < 0.05$).

Exhibit 5-3. Cumulative Three-Year Performance among 22 Primary Care Initiative Settings

		Unfavorable			No effect	Favorable		
		p<0.05	0.05<p<0.10	0.10<p<0.20	p>0.20	0.10<p<0.20	0.05<p<0.10	p<0.05
Core Outcomes								
Initiative	Setting	Medical Expenditures per Beneficiary	Outpatient ED Visits per 1,000 Beneficiaries	Hospital Admissions per 1,000 Beneficiaries	30-day Readmissions per 1,000 Beneficiaries			
CPC	Arkansas	\$36.00 ^d	4.0 ^d	-3.0 ^d	6.0 ^d			
	Colorado	-\$72.00 ^d	-6.0 ^d	2.0 ^d	-11.0 ^g			
	New Jersey	-\$228.00 ^e	0.0 ^d	-6.0 ^d	-1.0 ^d			
	New York	-\$312.00 ^e	3.0 ^d	-20.0 ^g	-5.0 ^d			
	Ohio/Kentucky	\$408.00 ^c	-11.0 ^d	14.0 ^d	-2.0 ^d			
	Oklahoma	-\$264.00 ^g	-22.0 ^g	-5.0 ^d	-5.0 ^d			
	Oregon	-\$228.00 ^f	-25.0 ^f	-10.0 ^e	-1.0 ^d			
FQHC-APCP		\$126.35 ^b	26.5 ^a	5.7 ^a	-3.9 ^d			
HCIA-PCR	CareFirst	\$130.67 ^d	-1.2 ^d	6.6 ^c	5.7 ^a			
	Cheyenne	N/A	N/A	N/A	N/A			
	Denver Health	-\$37.33 ^d	107.3 ^a	12.2 ^d	0.6 ^d			
	FLHSA	\$135.41 ^d	-10.8 ^d	4.1 ^d	-1.3 ^d			
	PBGH	N/A	N/A	N/A	N/A			
	PeaceHealth	-\$449.00 ^d	57.3 ^b	-35.3 ^g	-0.1 ^d			
	Sanford Health	\$205.50 ^b	-24.3 ^g	4.1 ^d	-1.5 ^d			
TransforMED	-\$339.20 ^g	-25.2 ^g	-10.3 ^f	-3.4 ^g				
MAPCP	Maine	\$673.20 ^a	-39.7 ^e	22.2 ^a	4.6 ^d			
	Michigan	-\$248.16 ^d	23.0 ^a	-5.5 ^d	-9.7 ^d			
	Minnesota	\$408.60 ^b	26.4 ^c	3.9 ^d	-18.9 ^e			
	New York	-\$116.04 ^d	11.0 ^d	-14.8 ^f	-11.3 ^d			
	North Carolina	\$241.56 ^d	-4.8 ^d	12.6 ^a	8.1 ^c			
	Pennsylvania	-\$311.04 ^d	-18.1 ^e	4.4 ^d	-0.9 ^d			
	Rhode Island	-\$17.04 ^d	-16.4 ^d	17.2 ^d	26.7 ^b			
Vermont	-\$324.84 ^g	42.6 ^a	4.6 ^d	-0.8 ^d				
Grand mean		-\$25.93	1.1	0.3	-0.9			
Notes: ED=emergency department; N/A = not applicable; PCMH= patient-centered medical home a: unfavorable, p<0.05; b: unfavorable, 0.05<p<0.10; c: unfavorable, 0.10<p<0.20; d: p>0.20 e: favorable, 0.10<p<0.20; f: favorable, 0.05<p<0.10; g: favorable, p<0.05								

HCIA TransforMED, a health IT intervention, is the only setting to show improvement across all four core outcomes at the $p < 0.10$ level. Two CPC settings (OK and OR) each showed improvement for at least two of the four outcomes at $p < 0.10$. In CPC OK, CPC OR, and HCIA-TransforMed, we observe the relationship between utilization and expenditures, where significant reductions in utilization appear to have led to reductions in expenditures. Significant reductions in hospital admissions or readmissions for CPC-CO, CPC-NY, MAPCP NY, HCIA-PeaceHealth, and MAPCP-PA with non-significant reductions in expenditures suggest that significant reductions in expenditures may be realized in the future.

In addition to these impacts from participation in the initiatives, some evaluators also tested the impacts of the PCMH model relative to a comparison group of non-PCMH practices. For example, the comparison group for evaluating the FQHC initiative includes some FQHCs that had achieved medical home recognition and some FQHCs that received technical assistance similar to much of the TA provided to initiative participants. Both of these confounding issues likely contribute to the lack of favorable impacts on the four core outcomes for Medicare beneficiaries attributed to the FQHCs participating in the initiative. The evaluator (RAND) also compared all FQHCs that had achieved Level-3 PCMH recognition (whether they were part of the treatment group or the comparison group) to several alternative comparison groups that either had not achieved Level-3 recognition or had not achieved any PCMH recognition. In these alternative analyses that focused on the effect of the PCMH model rather than the effect of participation in the CMS initiative, RAND found significant favorable PCMH impacts on utilization and expenditures (Kahn et al., 2016).

5.3 Summary and Discussion

When considered in the aggregate, the four initiatives included in our meta-analysis failed to significantly reduce cumulative expenditures or utilization as measured by the four core outcomes. Moreover, no initiative performed significantly better or worse than others. However, at the more granular level, two-fifths of settings within initiatives achieved significant improvement in at least one core outcome at $p < 0.10$, and three settings achieved significant reductions in both the total cost of care and at least one core utilization outcome at $p < 0.10$. For the most part, however, there was little correspondence across outcomes within settings. That is, in most settings, the impact of participating in an initiative did not produce consistent results relative to comparison groups across the four core outcomes. Over time (Years 1–3), performance within settings was generally consistent and did not tend to improve.

Because these outcomes are largely uncorrelated, it is likely that different setting features are associated with performance on each outcome and the factors associated with success on each outcome should be analyzed separately. We present meta-regression analyses examining potential reasons for these differences in Chapter 7.

6. Using Pooled Regression Analysis to Examine Subgroups

Primary care initiatives may have more (or less) favorable impacts on particular subgroups of Medicare beneficiaries, on practices with certain characteristics, or in health care markets with certain characteristics. This meta-evaluation gave us the opportunity to understand which populations and which practices experienced the largest impact from primary care initiatives. We were able to do this by combining beneficiary-level claims data across initiatives to determine the differential effects of primary care initiatives on the four core outcomes (medical expenditures without fees, outpatient ED visits, hospital admissions, and 30-day readmissions). We then used regression analysis with the pooled data to estimate the impacts of the primary care initiatives on beneficiary and practice subgroups.

As in the previous chapter, we included four of the six CMMI initiatives, for which we had patient-level analysis files reflecting two or three years of demonstration exposure. These include eight MAPCP state settings (NY, RI, VT, NC, MN, ME, MI, PA), seven CPC regional settings (AR, CO, NJ, NY, OH/KY, OK, OR), six of eight HCIA-PCR awardee settings (CareFirst, Denver Health, Finger Lakes Health Systems Agency, Peace Health, Sanford Health, and TransforMED), and FQHC.

The regression analysis of pooled beneficiary-level data enables us to address the key study question: Were there any common subgroups (e.g., beneficiaries with higher HCC risk scores, dual eligibles) that experienced differential impacts? We first summarize the methodology and then present the baseline model that tests whether collectively the four initiatives had a significant impact on health care expenditures and utilization. We then compare the impacts on subgroups of beneficiaries and primary care practices.

6.1 Methodology

We provide a brief overview of the data and methodology that we used in the analyses of the pooled, beneficiary-level claims data. Details of the methodology are available in *Appendix B*, and full results tables are available in *Appendix C*.

6.1.1 Primary Care Initiatives

We obtained final claims analysis files from the evaluators for CPC, FQHC, MAPCP, and six of the HCIA demonstrations. We converted all analysis files to the beneficiary-year level; while the level of observation is annual, we divided Medicare expenditures by twelve to obtain an average per-beneficiary-per-month (PBPM) expenditure, as was standard in the evaluator reports. Some of the initiatives did not have data for all subgroups and characteristics that we examined; *Exhibit 6-1* displays which variables of interest were available in each of the initiatives.

Exhibit 6-1. Data Availability for the Pooled Regressions

Variables	CPC	FQHC	HCIA	MAPCP
Beneficiary characteristics	●	●	●	●
Practice characteristics	●	●	—	●
Beneficiary survey data	●	●	—	●
Practice survey data	●	—	—	●
County-level data	●	●	—	●

6.1.2 Baseline Model: The Overall Impact of the Primary Care Initiatives

We used pooled regression analyses to identify the average impacts of these four initiatives on the four core outcomes (baseline model) and the moderating impact of beneficiary-level characteristics (e.g., dual-eligibility), practice-level features (e.g., practices with six or fewer practitioners), survey characteristics (both beneficiary-level and provider-level), and county-level demographics and health care market variables.

The purpose of the baseline model is to provide a benchmark assessment of the aggregate impact of each of the included initiatives on the four core claims-based outcomes—total Medicare expenditures, inpatient admissions, ER visits not leading to hospitalization, and unplanned hospital readmissions. We used an annual-level differences-in-differences framework, controlling for geographic region, calendar year, and the set of common beneficiary characteristics available across each of the evaluator’s final claims analysis files—age (categorically defined: < 65, 65–74, 75–84, > 84), female, non-white, disability (originally qualified for Medicare due to disability), dually eligible for Medicare and Medicaid, and HCC risk score. *Exhibit 6-2* presents the descriptive statistics for the beneficiary-level covariates separately for the combined demonstration group and for the combined comparison group. Mean values for the four core outcomes are also presented for the demonstration group and for the comparison group.

6.1.3 Triple Difference-in-Differences Models

To test whether the primary care initiatives had more (or less) favorable effects on beneficiaries with certain characteristics (or beneficiaries assigned to practices with certain characteristics), we estimate a triple difference-in-differences model. From this triple DiD model, we report three estimates for each of the four outcomes. For example, with the population that originally qualified for Medicare due to disability, we report (1) the collective impact of the primary care initiatives on disabled beneficiaries, comparing disabled beneficiaries in the demonstration group to disabled beneficiaries who were in the comparison group; (2) the collective impact of the primary care initiatives on the non-disabled beneficiaries, comparing non-disabled beneficiaries in the demonstration group to non-disabled beneficiaries who were in the comparison group; and (3) the difference between (1) and (2).

Exhibit 6-2. Summary Statistics for Beneficiary Characteristics and Outcomes at Baseline

Measure	Demonstration Group	Comparison Group
Number of beneficiaries	1,493,803	2,637,313
Number of practices	2,641	3,946
Beneficiaries from the CPC initiative (%)	28.85	26.86
Beneficiaries from the FQHC initiative (%)	19.58	16.48
Beneficiaries from the HCIA initiative (%)	10.37	10.10
Beneficiaries from the MAPCP initiative (%)	41.19	46.56
Beneficiaries under the age of 65 (%)	25.59	25.64
Beneficiaries aged 65 to 74 (%)	39.42	40.18
Beneficiaries aged 75 to 84 (%)	24.37	24.05
Beneficiaries over the age of 84 (%)	10.62	10.13
Female beneficiaries (%)	57.52	57.39
Non-White beneficiaries (%)	14.04	14.64
Disabled beneficiaries (%)	31.20	31.51
Dually eligible beneficiaries (%)	23.76	25.33
HCC risk score	1.02	1.02
Total Medicare Expenditures (\$)	828.11	811.34
All-cause inpatient admissions (per 1,000 beneficiaries)	264.53	262.68
ER visits not leading to hospitalization (per 1,000 beneficiaries)	534.93	568.21
% with any readmissions within 30 days of a discharge	10.60	10.42

6.2 Interpretation of the Baseline Model Results

Exhibit 6-3 presents the results of the baseline model for the four core outcomes using the pooled data. The estimates reported are the weighted average of each individual year estimate, where the weights are the relative number of demonstration beneficiaries present in each year. The estimates measure the impact of the four initiatives collectively and do not indicate whether any individual initiative had an impact on these outcomes. This analysis provides context for our subsequent testing of whether various subgroups experienced more favorable impacts than other subgroups, or than the pooled population as a whole. Across the four initiatives, the impact observed on the core outcomes, with the exception of ER visits, is in the theorized direction. However, none of the results are statistically significant at the 10 percent level.

Exhibit 6-3. The Collective Impact of the CPC, FQHC, HCIA, and MAPCP Initiatives on the Four Core Claims-Based Outcomes Using Pooled Data

Outcome	Estimate	90% Confidence Interval
Total Medicare expenditures (per beneficiary per month)	-11.42	[-25.85, 3.00]
All-cause inpatient admissions (per 1,000 beneficiary years)	-1.86	[-6.74, 3.02]
ER visits not leading to hospitalization (per 1,000 beneficiary years)	3.41	[-7.68, 14.51]
Any readmissions within 30 days of a discharge	-0.12	[-0.39, 0.16]

NOTES:

* $p < 0.10$

All estimates were produced using weighted least-squares.

Total Medicare expenditure estimates are interpreted as the difference in expenditure per beneficiary per month.

All-cause inpatient admission and ER visit estimates are interpreted as the difference in utilization per 1,000 beneficiary years.

Any readmission estimates are interpreted as the difference in probability of having any readmissions within 30 days of a live discharge during the past year.

6.3 Beneficiary Characteristics

Because many of the initiatives focused on care management and care coordination activities for beneficiaries with chronic conditions, we might expect to see more favorable impacts of the primary care initiatives on beneficiaries with more health care problems and higher health care needs. These beneficiaries might benefit more from the greater care coordination and other primary care initiatives aimed at providing more efficient care. For example, in the evaluation of the MAPCP Demonstration in Michigan, beneficiaries with multiple chronic conditions and those who were dually eligible experienced significantly slower growth in total Medicare expenditures, and the dually eligible beneficiaries also experienced decreases in all-cause admission rates. In the New York MAPCP Demonstration, beneficiaries with multiple chronic conditions had decreased rates of all-cause admissions.

We identified three consistently measured beneficiary characteristics across the analysis files for CPC, FQHC, MAPCP, and six of the HCIA demonstrations that identified subgroups of beneficiaries who might particularly benefit from greater care coordination. We tested whether these initiatives collectively had a significant impact on beneficiaries with greater health needs (i.e., those who had HCC scores in the top quartile in the baseline period, those who originally qualified for Medicare due to disability, and those who were dually eligible for Medicare and Medicaid) and whether the collective impact of the four initiatives was significantly larger or smaller for these beneficiaries compared to beneficiaries who did not have those characteristics. *Exhibit 6-4* presents our findings; only subgroups/outcomes with significant results are reported.

Exhibit 6-4. The Impact of the Initiative on Subgroups of Beneficiaries

Characteristic	Total Medicare Expenditures (PBPM)	Inpatient Admissions	ER Visits	Readmissions
HCC risk scores				
Beneficiaries in the upper quartile (range: 1.217-15.658)	-49.39*	ns	ns	ns
Beneficiaries in the lower three quartiles (range: 0.12-1.2169)	0.91	ns	ns	ns
Difference in effects	-50.31*	ns	ns	ns
Qualified for Medicare due to disability				
Disabled beneficiaries	-38.60*	ns	ns	0.44
Non-disabled beneficiaries	1.41	ns	ns	-0.43*
Difference in effects	-40.01*	ns	ns	0.87*
Dual-eligibility				
Dually eligible beneficiaries	ns	ns	ns	0.53
Non-dually eligible beneficiaries	ns	ns	ns	-0.42*
Difference in effects	ns	ns	ns	0.94*

NOTES:

* $p < 0.10$

All estimates were produced using weighted least-squares.

Total Medicare expenditure estimates are interpreted as the difference in expenditure per beneficiary per month.

All-cause inpatient admission and ER visit estimates are interpreted as the difference in utilization per 1,000 beneficiary years.

Any readmission estimates are interpreted as the difference in probability of having any readmissions within 30 days of a live discharge during the past year.

Specifically, we found the following:

- Among beneficiaries in the highest quartile of HCC scores, the initiatives collectively decreased the growth in **total Medicare expenditures** by \$49.39 PBPM. This estimate was statistically significantly different from the estimated impact among beneficiaries in all other quartiles of the distribution of HCC scores.
- Among beneficiaries who originally qualified for Medicare due to disability, the initiatives collectively decreased the overall growth in **total Medicare expenditures** by \$38.60 PBPM. The overall estimate was statistically significantly different from

the estimated impact among beneficiaries who did not originally qualify for Medicare due to disability.

- Among beneficiaries who did not qualify for Medicare due to disability, the initiatives collectively decreased the overall probability of having **any readmission within 30 days of a hospital discharge** by 0.43 percentage points. This estimate was statistically significantly different from the estimated impact among dually eligible beneficiaries.
- Among non-dually eligible beneficiaries, the initiatives collectively decreased the overall probability of having **any readmission within 30 days of a hospital discharge** by 0.42 percentage points. This overall estimate was statistically significantly different from the estimated impact among dually eligible beneficiaries.

The slower growth in Medicare expenditures for the beneficiaries in the highest quartile of HCC scores and for beneficiaries who originally qualified for Medicare through disability is consistent with our prior expectation that these initiatives could have a stronger impact on beneficiaries with high health care needs. We are uncertain as to why these initiatives are associated with lower probabilities of hospital readmission for non-dually eligible beneficiaries and non-disabled beneficiaries relative to dual eligibles and the disabled, respectively. It should be noted that the change in probability of 0.42–0.43 percentage points, while statistically significant, is not very large relative to the average of 10.6 percent of hospital stays among the combined demonstration groups with at least one readmission. Note also that there is significant overlap between the disabled and dually eligible beneficiaries: more than two thirds of the dually eligible beneficiaries originally qualified for Medicare due to disability. The model testing the differential effect on disabled beneficiaries includes a control for those who are dually eligible, and vice versa.

6.4 Practice-Level Characteristics

We also investigated whether the collective impacts across initiatives varied with respect to several practice-level characteristics. The practice-level characteristics that were available for each initiative depended on the data elements that were collected by each evaluator. We present findings separately for three sets of measures: (1) those that were available on the analytic claims-based files, (2) those that we created from patient-level surveys capturing patient care experience, and (3) those that we created from practice level surveys capturing areas of practice transformation.

The CPC, MAPCP, and FQHC evaluators collected two practice-level measures and included them in their analytic claims-based files: (1) whether the practice was a multispecialty practice, and (2) the number of practitioners at each practice. We used the latter measure to create a dichotomous variable indicating whether the practice had fewer than six practitioners versus six or more practitioners. Comparing across these practice characteristics, we then contrasted CPC, FQHC, and MAPCP Demonstration practices with comparison practices from the evaluations of each of these initiatives. We found that both practice-level measures that were included in the evaluators' analytic claims-based files were associated with statistically significant and favorable outcomes (see *Exhibit 6-5*).

Exhibit 6-5. The Impact of the Initiatives on Subgroups Based on Practice Characteristics

Characteristic	Total Medicare Expenditures (PBPM)	Inpatient Admissions	ER Visits	Readmissions
Presence of specialty providers				
Practices with specialty providers	15.74	6.85	32.01*	ns
Practices without specialty providers	-11.32*	-6.35*	-21.34*	ns
Difference in effects	27.06*	13.20*	53.34*	ns
Practice size				
Fewer than six practitioners	-10.37	-6.90*	-12.01	ns
Six or more practitioners	8.00	3.39	16.58	ns
Difference in effects	18.37*	10.28*	28.59*	ns

NOTES:

* $p < 0.10$

All estimates were produced using weighted least-squares.

All models controlled for age, gender, race/ethnicity, disability, dual-eligibility, and HCC risk score, as well as the practice characteristics listed in the table.

Total Medicare expenditure estimates are interpreted as the difference in expenditure per beneficiary per month.

All-cause inpatient admission and ER visit estimates are interpreted as the difference in utilization per 1,000 beneficiary years.

Any readmission estimates are interpreted as the difference in probability of having any readmissions within 30 days of a live discharge during the past year.

Specifically, we found the following:

- Among practices with *no specialty providers*, participation in the initiatives collectively decreased
 - Growth in **total Medicare expenditures** by \$11.32 PBPM, relative to practices with no specialty providers in the comparison group. This estimate was also statistically significantly different from the overall estimated impact among multispecialty practices.
 - The rate of **all-cause inpatient admissions** by 6.35 admissions per 1,000 beneficiary years. This estimate was also statistically significantly different from the overall estimated impact among multispecialty practices.
 - The rate of **ER visits not leading to hospitalization** by 21.34 visits per 1,000 beneficiary years. This estimate was also statistically significantly different from the overall estimated impact among multispecialty practices.

- Among practices *with fewer than six practitioners*, participation in the initiatives collectively decreased the rate of **all-cause inpatient admissions** by 6.90 admissions per 1,000 beneficiary years, relative to practices with fewer than six practitioners in the comparison group. This estimate was also statistically significantly different from the estimated overall impact among practices with six or more practitioners.

Using patient-level survey data, we also created five composite scores reflecting different aspects of patient experience with the care they received. Specifically, we created scores at the practice level for (1) timely access to appointments, care, and information; (2) how well providers communicate with patients; (3) attention to care from other providers; (4) provider supports patients in taking care of their own health; and (5) provider discusses medication decisions. We used regression modeling to compare practices with higher scores (i.e., \geq 75th percentile) versus practices with lower scores (i.e., $<$ 75th percentile), and to contrast CPC, FQHC, and MAPCP Demonstration practices with comparison practices from the evaluations of the CPC and FQHC initiatives. We found that few of the practice-level scores derived from patient surveys were associated with more statistically significant favorable outcomes (see *Exhibit 6-6*).

Exhibit 6-6. The Impact of the Initiatives on Subgroups Based on Practice-Level Patient Experience Scores

Patient Experience Score	Total Medicare Expenditures (PBPM)	Inpatient Admissions	ER Visits	Readmissions
Getting timely appointments, care, and information				
Practices with scores $<$ 75th percentile	ns	ns	ns	ns
Practices with scores \geq 75th percentile	ns	ns	ns	ns
Difference in effects	ns	ns	ns	ns
How well providers communicate with patients				
Practices with scores $<$ 75th percentile	ns	ns	ns	ns
Practices with scores \geq 75th percentile	ns	ns	ns	ns
Difference in effects	ns	ns	ns	ns
Attention to care from other providers				
Practices with scores $<$ 75th percentile	ns	ns	ns	-0.45
Practices with scores \geq 75th percentile	ns	ns	ns	0.54
Difference in effects	ns	ns	ns	1.00*

(continued)

Exhibit 6-6. The Impact of the Initiatives on Subgroups Based on Practice-Level Patient Experience Scores (continued)

Patient Experience Score	Total Medicare Expenditures (PBPM)	Inpatient Admissions	ER Visits	Readmissions
Provider supports patients in taking care of their own health				
Practices with scores < 75th percentile	ns	ns	ns	-0.042
Practices with scores ≥ 75th percentile	ns	ns	ns	-0.83*
Difference in effects	ns	ns	ns	-0.78
Providers discuss medication decision scale				
Practices with scores < 75th percentile	ns	ns	ns	ns
Practices with scores ≥ 75th percentile	ns	ns	ns	ns
Difference in effects	ns	ns	ns	ns

NOTES:

* $p < 0.10$

All estimates were produced using weighted least-squares.

All models controlled for age, gender, race/ethnicity, disability, dual-eligibility, and HCC risk score, as well as the practice characteristics listed in the table.

Total Medicare expenditure estimates are interpreted as the difference in expenditure per beneficiary per month.

All-cause inpatient admission and ER visit estimates are interpreted as the difference in utilization per 1,000 beneficiary years.

Any readmission estimates are interpreted as the difference in probability of having any readmissions within 30 days of a live discharge during the past year.

However, as shown in *Exhibit 6-6* we found a few statistically significant differences for the readmissions outcome.

Using practice-level survey data, we also created five composite scores reflecting aspects of primary care transformation. Specifically, we created composite scores for (1) access to care, (2) patient-centered interactions, (3) care coordination, (4) care management, and (5) quality improvement (QI) activities. We used regression modeling to compare practices with higher scores (i.e., ≥ 75th percentile) versus practices with lower scores (i.e., < 75th percentile), and to contrast CPC and MAPCP Demonstration practices with comparison practices from the evaluations of the CPC initiative. We found that several primary care transformation scores were associated with more favorable outcomes in the CPC and MAPCP initiatives (see *Exhibit 6-7*).

Exhibit 6-7. The Impact of the Initiatives on Subgroups Based on Practice-Level Primary Care Transformation Scores

Primary Care Transformation Score	Total Medicare Expenditures (PBPM)	Inpatient Admissions	ER Visits	Readmissions
Access to care				
Practices with scores < 75th percentile	ns	-4.58	ns	-1.06*
Practices with scores ≥ 75th percentile	ns	-13.62*	ns	-0.34
Difference in effects	ns	-9.04	ns	0.72
Patient-centered interactions				
Practices with scores < 75th percentile	ns	-9.23*	ns	-0.77*
Practices with scores ≥ 75th percentile	ns	-2.04	ns	-0.89*
Difference in effects	ns	7.19	ns	-0.12
Care coordination				
Practices with scores < 75th percentile	ns	-0.01	ns	-0.32
Practices with scores ≥ 75th percentile	ns	-16.83*	ns	-1.71*
Difference in effects	ns	-16.82*	ns	-1.39*
Care management				
Practices with scores < 75th percentile	ns	ns	ns	-0.99*
Practices with scores ≥ 75th percentile	ns	ns	ns	-0.59
Difference in effects	ns	ns	ns	0.40
Quality Improvement activities				
Practices with scores < 75th percentile	4.45	1.98	ns	-0.68
Practices with scores ≥ 75th percentile	-30.09*	-16.75*	ns	-0.76*
Difference in effects	-34.55*	-18.73*	ns	-0.081

NOTES:

* $p < 0.10$

All estimates were produced using weighted least-squares.

All models controlled for age, gender, race/ethnicity, disability, dual-eligibility, and HCC risk score, as well as the practice characteristics listed in the table.

Total Medicare expenditure estimates are interpreted as the difference in expenditure per beneficiary per month.

All-cause inpatient admission and ER visit estimates are interpreted as the difference in utilization per 1,000 beneficiary years.

Any readmission estimates are interpreted as the difference in probability of having any readmissions within 30 days of a live discharge during the past year.

Specifically, we found that:

- Among practices with higher self-reported scores (i.e., \geq 75th percentile) for care coordination, participation in CPC or MAPCP decreased
 - The rate of **all-cause inpatient admissions** by 16.83 admissions per 1,000 beneficiary years, relative to practices with higher self-reported scores for care coordination in the comparison group. This estimate was statistically significantly different from the estimated impact among practices with lower self-reported scores (i.e., $<$ 75th percentile) for care coordination.
 - The probability of **any readmissions within 30 days of a discharge** by 1.71 percentage points. This estimate was statistically significantly different from the estimated overall impact among practices with lower self-reported scores (i.e., $<$ 75th percentile) for care coordination.
- Among practices with higher self-reported scores (i.e., \geq 75th percentile) for quality improvement (QI) activities, participation in CPC or MAPCP decreased
 - Growth in **total Medicare expenditures** by \$30.09 PBPM, relative to practices with higher self-reported scores for QI activities in the comparison group. This estimate was statistically significantly different from the estimated impact among practices with lower self-reported scores (i.e., $<$ 75th percentile) for QI activities.
 - The rate of **all-cause inpatient admissions** by 16.75 admissions per 1,000 beneficiary years. This estimate was statistically significantly different from the estimated overall impact among practices with lower self-reported scores (i.e., $<$ 75th percentile) for QI activities.

6.5 County-Level Characteristics

For CPC, MAPCP, and most FQHC practices, we could identify the county where the practice was located. For these practices, we derived and merged the following county-level characteristics: (1) the number of Medicare beneficiaries per capita, (2) whether the county was classified as rural or urban, (3) whether the county had one or more areas designated as a primary care shortage area, (4) the percentage of Medicare beneficiaries who were dually eligible, (5) the Medicare Advantage penetration rate, (6) the number of persons aged 65 or older per capita, (7) the number of hospital beds per capita, (8) the rate of ED (outpatient and inpatient) visits, and (9) whether the county was designated as a high poverty county. Comparing across county characteristics, we then contrasted CPC, FQHC, and MAPCP Demonstration practices with comparison practices from the evaluations of each of these initiatives. We found virtually no evidence that initiative impacts were associated with any of these county-level characteristics.

6.6 Summary

We focused these pooled regression analyses on whether the initiatives' impacts differed significantly for population or practice subgroups. Of greatest importance, we found that beneficiaries originally eligible for Medicare due to disability and beneficiaries with the highest

quartile of baseline HCC scores experienced slower growth in Medicare expenditures under the initiatives.

We also found impacts on different types of practices. We found there was significantly slower growth in Medicare expenditures and lower rates of inpatient admissions and ER visits among both practices with fewer than six practitioners and those that were primary care-only practices, relative to practices with these same characteristics who did not participate in the initiatives.

We also identified practices that scored highly on five composite measures from the overlapping elements of the practice surveys, limited only to CPC and MAPCP practices and CPC comparison practices with responses to the survey. We found that practices with high care management scores and high QI scores saw slower growth in Medicare expenditures and lower rates of inpatient admissions. We also found that practices with high care coordination scores saw lower rates of inpatient admissions.

One important limitation to these analyses is that, within each of the separate evaluations, the comparison groups were chosen and propensity-score weighted to achieve balance with the covariates for the entire demonstration group, not necessarily subgroups of the demonstration group. There may be important differences between the demonstration and comparison groups in our analysis, particularly for the smaller subgroups of beneficiaries or practices. Nonetheless, we believe the pooled regression results add value by bringing together data from multiple evaluators to examine these subgroups with more precision than any individual evaluation could on its own.

7. Meta-regression and QCA Analyses of Features that Contribute to High or Low Performance on Utilization and Expenditures

To identify which features contributed to high performance, we conducted both meta-regression and a Qualitative Comparative Analysis (QCA). In this section, we briefly summarize the findings from these two approaches.

7.1 Meta-regression: Identifying Constellations of Features Associated with Performance on the Four Core Outcomes

As discussed in Chapter 5, we observed significant heterogeneity in setting performance in each of the four core outcomes. This indicates there was more variation in setting performance than would be expected from sampling theory. In Chapter 5, we also observed limited correspondence of performance among the four core outcomes, with the exception of the relationship between hospital admissions and medical expenditures. Knowing a setting's performance on one outcome gave little insight as to how well it performed on other outcomes. Therefore, we examined first the bivariate relationship between each of the factors and features documented using structured coding of evaluator reports and responses received from evaluators in the Evaluator Summary Form (ESF) with each of the four core outcomes (expenditures without fees, outpatient ED visits, hospital admissions, and 30-day hospital readmissions). These meta-correlations are presented in *Appendix E, Exhibit E-3*. We then used meta-regression modeling to identify the constellation of features systematically associated with performance on each of the four core outcomes, and the results of this meta-regression analysis are described below. Due to the varying goals of the initiatives, the differences in how each setting sought to achieve those goals, and to ensure all settings contribute to the analysis, all findings are based on random effects models.

In the analyses that follow, the coefficient presented in each random-effects meta-regression model represents the systematic difference in setting-level difference-in-difference (DiD) effects associated with a one-unit increase in the feature (see *Appendix E, Exhibit E-1* for features' measurement properties and *Exhibit E-2* for feature means and frequencies). In other words, the coefficients in these models estimate the systematic contribution of each feature to the DiD performance estimates achieved by the 22 settings providing sufficient data for these analyses (eight MAPCP states, seven CPC regions, six HCIA awards, and FQHC). Most features were coded as binary attributes, for example, receiving TA from a particular source versus not receiving TA from that source. In the meta-regression models, the coefficient reflects the unstandardized relationship between the feature and the outcome after controlling for the contribution of other features in the model.

To better understand how setting factors and features are associated with setting performance on the four core measures, we used random-effects stepwise meta-regression to identify those factors and features that best explained differences in outcomes for the four core measures. The unstandardized coefficients in these models describe the systematic (linear) increase or decrease in the outcome across the settings when all other features of the model are accounted for.

While we believed that all the features selected to be tested in the meta-regression had capacity to influence outcomes, the stepwise regression approach is empirical. It starts with entering the feature that explains the greatest amount of variation in findings, then each remaining feature is tested to identify which feature explains the greatest amount of the

remaining variation. That feature is retained in the model, and all remaining features are again tested to identify the feature with the greatest explanatory value. This continues until the variables entered no longer improve the fit of the model or when all variation attributable to between-study differences is explained, whichever comes first. Stepwise meta-regression optimizes the model by retaining successively the features that explain the greatest amount of the variation given features already in the model. The stepwise approach was used in this context because of the need to identify a small number of features with the most explanatory power for a given outcome in a situation with only a small number of settings.

On the basis of this method, the following models emerged as most predictive of the four core outcomes (see *Appendix E, Exhibits E-4 through E-8* for full models and fit statistics). It should be noted that meta-analysts distinguish between the variability among estimates expected from and attributable to sampling error and the variation attributable to between-study differences (referred to as heterogeneity). The percentage of the total variation attributable to between-study differences is estimated by the I^2 statistic and is the variability in results available to be explained in meta-regression models. Thus, in these models, the R^2 -adjusted value is the estimated percentage of between-study variation (I^2) explained by the models—not the percentage of total variation in the distribution of setting results.

Predicting effects on expenditures without fees, net of hospital admissions. Because hospital admissions were highly correlated with expenditures and are a significant driver of expenditures, we elected to treat admissions as a confounding variable and to statistically control for its influence before assessing the impact of other setting features on expenditures (Lipsey, 1990). Controlling for confounding variables increases the ability of a model to detect other substantive sources of variance.

As anticipated, the meta-regression results indicate that for every additional **hospital admission** per 1,000 beneficiaries per year, medical expenditures increased by \$15.20 per enrollee per year ($p < 0.01$) relative to comparators. After accounting for the additional costs due to hospital admissions, the meta-regression indicates that **settings that used data reports to change activities** decreased expenditures by \$135.00 per enrollee per year ($p = 0.04$), while settings that **received TA from CMS or a CMS subcontractor** reduced expenditures by \$153.70 per enrollee per year ($p = 0.05$) relative to comparators. With 53.7 percent of the initial variation in setting results attributable to between-setting differences, the three variables in this model explained 100 percent of the variation attributable to between-setting differences and are significant ($p < 0.01$), indicating an association of at least one of the covariates with the size of the treatment effect (see *Exhibit 7-1*).

Exhibit 7-1. Meta-regression Predicting Medical Expenditures

Features	Difference in Cumulative Year 3 Expenditures			
	Coef	SE	90% CI	p
TA provision: CMS or CMS sub	-153.73	73	(-281 to -27)	0.05
Used data report to change activities ¹	-134.95	61	(-241 to -29)	0.04
Hospital admissions	15.19	4	(8 to 22)	<0.01
Constant	452.55	191	(2 to 783)	0.03

Notes: CI = confidence interval; Coef = coefficient; p = p-value; SE = standard error.

1. Mean imputation by initiative setting was used to fill in missing values.

Predicting effects on outpatient ED visit. As observed in preliminary meta-correlation analyses, several beneficiary attributes were associated with setting impact on outpatient ED visits. The strongest of these was the percent of male Medicare beneficiaries in the treatment group covered by the setting. Given the strength of that relationship, we ran meta-regression models both with and without percent of males. Across settings, each 1-percent increase in the **percentage of male enrollees** was associated with 8.3 more outpatient ED visits per 1,000 enrollees per year ($p < 0.01$). After controlling for the percentage of males in the setting, **provision of peer-to-peer TA** was associated with 17.2 fewer outpatient ED visits per 1,000 enrollees per year ($p = 0.12$). With 78.0 percent of the total distribution of setting results attributable to between setting differences, the model explained 67.4 percent of the between-setting variability. The model was significant ($p < 0.01$), indicating an association of at least one of the covariates with the size of the treatment effect (see *Exhibit 7-2*).

**Exhibit 7-2. Meta-regression Predicting Outpatient ED Visits—
Model Including Percent Males**

Features	Difference in Cumulative Year 3 Outpatient ED visits			
	Model including % males			
	Coef	SE	90% CI	p
% male	8.3	1.8	(5.2 to 11.4)	<0.01
TA mode: Peer-to-peer	-17.2	10.5	(-35.3 to 0.9)	0.12
Constant	-334.9	73.7	(62.4 to -207.4)	<0.01

Notes: CI = confidence interval; Coef = coefficient; p = p-value; SE = standard error.

After omitting percent of male enrollees since it is not a model design feature, we ended up with a model which includes Medicare payment per practice, two measures of TA, state convener, and the safety net medical home composite. In this final model, we find that each \$10,000 increase in **Medicare payments per practice** was associated with an increase of 0.37 outpatient ED visits per 1,000 enrollees per year ($p < 0.01$). Settings that arranged for their **own TA or received TA from sources other than CMS or a CMS contractor** reduced ED visits by 47.5 per 1,000 enrollees relative to comparators ($p < 0.01$). **Direct assistance TA**, however, was associated with 21.0 more ED visits per 1,000 enrollees ($p < 0.01$), while having a state convener was associated with 38.8 more ED visits ($p < 0.01$). Finally, each one point improvement in our **safety net medical home composite score** (scaled 1-100), was associated with a decrease of 0.4 ED visits per year per 1,000 enrollees relative to comparators. The model was significant ($p < 0.01$), and explained 100 percent of the initial 78.0 percent variation attributable to between-setting differences (see *Exhibit 7-3*).

**Exhibit 7-3. Meta-regression Predicting Outpatient ED Visits—
Model Excluding Percent Males**

Features	Difference in Cumulative Year 3 Outpatient ED visits			
	Model excluding % males			
	Coef	SE	90% CI	p
Convener: State	38.8	7.3	(26.1 to 51.5)	<0.01
TA provision: other	-47.5	7.6	(-60.7 to -34.2)	<0.01
TA mode: Direct Assistance	20.9	7.4	(8.0 to 33.9)	0.01
Medicare payment per practice (in \$10k)	0.4	0.1	(0.2 to 0.5)	<0.01
Safety net medical home composite	-0.4	0.2	(-0.7 to -0.1)	0.05
Constant	23.5	13.4	(0.1 to 46.9)	0.10

Notes: CI = confidence interval; Coef = coefficient; p = p-value; SE = standard error.

Predicting effects on hospital admissions: Hospital admissions was another outcome affected by beneficiary characteristics; for this outcome, the **percent of dual enrollees** was associated with an increase of 0.3 hospital admissions per 1,000 enrollees per year (p=0.05), indicating that relative to comparators and accounting for other variables in the model, primary care transformation was less effective in reducing hospital admissions among settings with the highest percentage of dual enrollees. **Hiring new staff** was associated with 6.7 fewer hospital admissions per 1,000 enrollees per year (p<0.01), while settings with a larger **percentage of large practices** (practices with 6 or more practitioners) was associated with 0.4 fewer hospital admissions per 1,000 enrollees per year relative to comparators (p = 0.02). **Requiring or favoring PCMH accreditation** was associated with an increase of 16.2 hospital admissions per 1,000 beneficiaries per year (p<0.01); and each one point improvement in our **safety net medical home composite score** (scaled 1-100), was associated with an additional 0.5 hospital admissions per 1,000 beneficiaries per year (p<0.01). Overall, the model was significant (p<0.01) and explained 80.3 percent of the initial 59.1% variation attributable to between-setting differences (see *Exhibit 7-4*).

Exhibit 7-4. Meta-regression Predicting Hospital Admissions

Features	Difference in Cumulative Year 3 Hospital Admissions			
	Coef	SE	90% CI	p
% duals	0.3	0.1	(0.1 to 0.5)	0.05
% large practices	-0.4	0.1	(-0.7 to -0.1)	0.02
Required or favored accreditation	16.2	5.2	(7.0 to 25.3)	<0.01
Hired new staff ¹	-6.7	2.2	(-10.5 to -2.9)	<0.01
Safety net medical home composite	0.5	0.1	(0.2 to 0.7)	<0.01
Constant	-6.9	10.0	(4.4 to 10.6)	0.50

Notes: CI = confidence interval; Coef = coefficient; p = p-value; SE = standard error.

1. Mean imputation by initiative setting was used to fill in missing values.

Predicting effects on 30-day readmissions: One feature, direct assistance TA, was sufficient for explaining all 37.9 percent of the variation in results attributable to between setting differences in 30-day readmissions. Settings which provided **direct assistance TA** experienced 7.2 fewer 30-day readmissions per 1,000 enrollees per year ($p < 0.01$; see *Exhibit 7-5*).

Exhibit 7-5. Meta-regression Predicting 30-day Readmissions

Features	Difference in Cumulative Year 3 30-Day Readmissions			
	Coef	SE	90% CI	p
TA mode: Direct Assistance	-7.2	1.9	(-10.4 to -4.0)	<0.01
Constant	4.9	1.6	(2.2 to 7.7)	<0.01

Notes: CI = confidence interval; Coef = coefficient; p = p-value; SE = standard error.

7.1.1 Meta-regression Discussion and Conclusions

The meta-regression analyses identified multiple features associated with improved outcomes. Different features are included in the meta-regression model for each core outcome. As documented in Chapter 5, with the exception of hospital admissions' relationship with expenditures, setting performance varied on the four core outcomes. Not surprisingly then, the results of the meta-regression document that the setting features associated with each of the four core outcomes likewise varied. Only two features contribute to more than one model. Our safety net medical home composite score was associated with small reductions in ED visits and small increases in hospital admissions. Provision of TA was significantly associated with significant improvements in three of the four outcomes, although the mode or provider of TA varied by outcome; hospital admissions was the sole outcome that did not have any TA variable in its final meta-regression model. As noted previously, direct assistance TA was associated with increased outpatient ED visits and decreased 30-day hospital readmissions. The relationship between direct assistance TA and increased ED visits must be interpreted with caution, however, as it is possible in some cases that poorer initial performance on an outcome may have led to an increase in provision of direct assistance TA. TA from CMS or a CMS subcontractor was associated with decreased medical expenditures, while TA provided by the setting or a source other than CMS was associated with reduced ED visits.

7.2 Qualitative Comparative Analysis (QCA)

QCA assesses features in combination with each other to identify necessary and sufficient pathways to successful reduction in cost. For this analysis, we assessed whether the initiative settings achieved a cumulative reduction in expenditures. We defined *successful* reduction in cost, scored as 1.0, as having a favorable effect as demonstrated by statistically significant ($p < 0.10$) reduction in growth of expenditure among beneficiaries in demonstration practices relative to comparison group practices. Moderate reduction in cost, scored as 0.66, referred to a *reduction* in growth of expenditure among beneficiaries in demonstration practices relative to comparison group practices, but reduction in growth does not reach statistical significance. We scored settings as lacking successful reduction in cost at 0.33 when they demonstrated an *increase* in growth of expenditures among beneficiaries in demonstration practices relative to comparison group practices, but increase in growth does not reach statistical significance; settings that

showed a statistically significant increase in growth of expenditure among beneficiaries in demonstration practices relative to comparison group practices were scored as 0.

For the demonstration features, we drew on the Evaluation Summary Form (ESF) and abstraction from the evaluators’ annual reports to identify demonstration features to include in an analysis. The project team identified several features: (1) high level of leadership experience with implementing activities similar to those in the initiative, (2) having adequate funding to support the demonstration, (3) high level of practice participation in technical assistance (TA), (4) high level of practice usage of data feedback reports, and (5) having adequate levels of staffing to support demonstration activities in practices. Because of the small number of cases, we were limited in the number of features we could include in our analysis.

For the pathways to high and low performance, we conducted a conventional fuzzy set QCA in R SetMethods and QCA packages and employed best practices (Schneider & Wagemann, 2012). *Appendix F* describes the analytic steps. *Exhibit 7-6* displays the settings and the characteristics they share as well as the consistency value (a parameter of fit) for high performance. Scores under .8 consistency are not considered combinations for high performance.

Exhibit 7-6. Characteristics of the Settings and Parameter of Fit for Successful Reduction in Cost

Leadership	Financial adequacy	TA	Feedback	Staff Adequacy	# of Cases in Combination	Consistency	Cases
1	1	1	1	1	5	.736	CPC-AR CPC-NJ* CPC-NY* HCIA-Finger Lakes MAPCP-ME
1	1	0	1	1	3	.925	CPC-CO* CPC-OH CPC-OK*
0	1	0	0	1	1	.994	MAPCP-RI*
1	1	0	0	1	1	.977	MAPCP-MI*
1	0	0	1	1	1	.939	CPC-OR*
1	0	0	0	0	1	.799	FQHC
1	0	1	0	0	1	.796	MAPCP-PA*

1 = present at high level, 0 = not present at high level

In fuzzy sets, cases have simultaneous membership in more than one combination; the truth table shows the case membership in the row with the “best fit” or (highest membership value).

With five conditions, 32 combinations, we are only showing rows with cases; all other rows not shown lacked cases.

* Settings scored as achieving .66 or higher on reduction in expenditures.

7.2.1 Analysis of Higher Performance

We analyzed these features to determine which combinations of features consistently lead to high performance. *Exhibit 7-7* shows the two solution terms and the cases in each solution; for full detail on the solutions and parameters of fit, see *Appendix F*. For this model, we tested for and found no individual necessary conditions. Two additional combinations were highly sufficient for the outcome of reduced expenditures: (1) having a high level of financial adequacy AND having a high level of staff adequacy AND not having a high level of practice participation in TA or (2) having leadership experience AND having a high level of practice usage of feedback reports AND having a high level of staff adequacy AND not having a high level of practice participation in TA.

Exhibit 7-7. QCA Solutions and Settings with This Combination of Features

Solution	Cases in this solution
Having a high level of financial adequacy AND having a high level of staff adequacy AND not having a high level of practice participation in TA	MAPCP-RI* MAPCP-MI* CPC-CO* CPC-OH CPC-OK*
Having leadership experience AND having a high level of practice usage of feedback reports AND having a high level of staff adequacy AND not having a high level of practice participation in TA	CPC-OR* CPC-CO* CPC-OH CPC-OK*

* Setting scored as achieving .66 or higher on reduction in expenditures.

The two solution terms point to the importance of initial investment in practice transformation resources to achieve an ultimate reduction in costs. For the first solution term (having financial adequacy AND having staff adequacy AND not having a high level of practice participation in TA), settings received financial support through the initiatives, which enabled them to hire additional staff, such as care managers, care coordinators, nurses, and advanced practice providers. Thus, the financial support and staff adequacy went hand-in-hand.

In the second solution term (having leadership experience AND having a high level of practice usage of feedback reports AND having a high level of staff adequacy AND not having a high level of practice participation in TA), leaders who have implemented similar efforts have the expertise to establish new, but similar, initiatives or could readily expand upon extant ones. These leaders may have processes, protocols, partners, or other resources that they drew upon to achieve initiative goals; thus, this experience may have increased implementation efficiency. Further, experienced leaders understand practices and how practices use data—whether from a formal feedback report or from the practice EHR. This knowledge allows such leaders to craft better strategies and TA for engaging practices in their use of data feedback reports. For example, this solution includes the CPC settings; the convener, CMS had experience with poor uptake of data feedback reports in other settings. Thus, they built in review of feedback reports as a milestone for practices.

Interestingly, not having a high level of practice participation in TA was a common feature in both solution terms. Although this is a counterintuitive finding, the initiatives in these

solutions frequently represented high capacity practices that may not depend on intensive TA to support practice transformation and, ultimately, reductions in expenditures. For example, MAPCP RI and MI reported practices had mixed opinions of the value of the technical assistance offered. Practices reported the technical assistance offered was not useful for advanced practices, took too much time away from other practice responsibilities, and the number of technical assistance opportunities exceeded the availability of staff to attend (Nichols et al., 2017). Although CPC generally included practices that had already achieved a level of NCQA recognition or had experience with transforming into PCMHs, CPC required practices to attend a minimum number of TA sessions (Peikes, Ghosh, et al., 2016). However, more advanced CPC practices may not depend on TA, particularly individual practice assistance, compared to practices in other settings struggling to meet CPC milestones (Peikes, Anglin, et al., 2016).

7.2.2 Analysis of Increases in Expenditures (Lower Performance)

We conducted a QCA for increases in expenditures performance, but we found no combinations that led to increases in expenditures at a strong level.

7.2.3 Key Lessons from QCA for High Performance

Although the QCA shows how the setting features work together to yield high performance, looking across the high-performing settings and their experience reveals several key insights:

- Adequate financial support can enable settings to hire staff needed to succeed in reducing Medicare expenditures.
- Having experienced leadership enabled settings to draw on extant processes, relationships, and other less tangible resources to achieve a reduction in costs (when combined with efforts to ensure higher uptake of data feedback reports).
- The QCA finding that settings that achieved savings on the Medicare cost of care did not have a high level of practice participation in TA is consistent with the MAPCP evaluation finding that practices reported mixed views on the TA that was offered in their states, including comments from more advanced practices that the TA provided was not useful for them because it was not tailored for their higher level of performance and occupied too much staff time. This underscores the importance of tailoring TA to account for different levels of practice transformation within an initiative or setting.

7.2.4 Limitations to Meta-Analysis and QCA

Both the meta-regression and the QCA analyses are limited by relatively small numbers of observations (22 settings included in the meta-regression and 13 settings in the QCA). Also, the ESF ratings are based on evaluators' perceptions at the setting level, and the abstractions rely on what the evaluators reported, which was not consistent across all evaluators and all reports. Further, the level of financial and staff adequacy may not have been uniform across all practices within a setting; rather the evaluators based their assessment of adequacy on most practices in the setting. Finally, we used the evaluator reports to explain the QCA findings, but often the reports contain initiative-level syntheses, rather than setting-level information. Because most

information is reported at the initiative level for CPC, we could not fully disentangle differences across CPC settings and why some reduced expenditures while others did not.

7.3 Summary and Discussion

Meta-regression and QCA are different methods for identifying features associated with initiative and setting success. Based as they are on different assumptions and purposes, the results from these two approaches are unlikely to produce similar findings. Meta-regression attributes variation in outcomes uniquely to variables and, if independent variables co-occur, will attribute variation to only one of the correlated variables. The other variable is omitted from the model as it does not identify unique variation. QCA is fundamentally different in that it assesses combinations of conditions that are necessary or sufficient for an outcome.

The QCA and meta-regression analyses also differ in how the dependent variable is constructed. The QCA focused on one core outcome (Medicare expenditures before fees), while the meta-regression considered each of the four core outcomes (independently). This reflects the different traditions each approach acknowledges and the standards that each approach must meet. These approaches are complementary, and the results obtained from each provide insights into how settings achieved success. We reached three key conclusions from these two analyses:

- Provision and receipt of TA can promote better performance, but practices' interest in TA and the benefit they perceive depends in part on how well the TA is tailored for differences in practices' level of transformation to becoming an advanced primary care practice.
- Feedback reports can improve performance but only to the extent practices actually use the data, which varied widely across initiatives.
- Adequate financial support allowed the hiring of new staff and likely reinforced practices' ability to transform.

This page has been left blank for double-sided copying.

8. Evaluation Challenges

This section addresses the most important challenges encountered by CMS and the evaluators in evaluating the initiatives. Our findings are based on the evaluation reports and interviews conducted for this study with CMMI evaluation project officers and the evaluators of each initiative.

8.1 Length of Time for Transformation and Evaluation of Impacts

Key informants interviewed for this study distinguished two issues related to the length of the evaluations: (1) the length of time it took for primary care practices to transform to advanced primary care practices and (2) the length of time it took for the impacts of the primary care delivery changes to be observed.

Almost all key informants thought that the time periods established in each initiative for practices to transform were sufficient. Even though transformation activities can take multiple years, most informants judged that practices had sufficient time to make the changes required by each initiative. However, the pace of change differed by initiative, which had implications for the evaluations. In CPC, CMMI set periodic milestones, and transformation activities were spread out over a multi-year period. In the FQHC initiative, the changes were more compressed, as almost half of the FQHCs that achieved NCQA Level-3 PCMH recognition did so in the final quarter of the three-year initiative time frame.

The pace of transformation also varied for certain types of practices. Small, inexperienced, and under-resourced practices had more difficulty transforming, especially in adopting or adapting health IT to meet new requirements and in hiring new staff whose roles then had to be integrated into the practice workflow. These challenges often led to delays in implementation, which increased the length of time it took for practices to meet initiative milestones or goals and, in turn, reduced the length of time in which the impacts of an initiative could be observed (exposure). Evaluators noted that this variation in exposure was also a challenge.

On the other hand, many of the key informants judged that the time periods to observe the impacts of transformation were too short. Part of this stemmed from the fact that some practices accomplished many of their transformation changes near the end of the initiative, as in the FQHC demonstration. As a result, these changes may have come too late to have an impact on the four core outcomes. A second factor is that certain transformation activities have impacts that can take longer to observe. For example, care coordination activities might be expected to affect the total cost of care more quickly than improving care for diabetics—some of whose major complications in the absence of the intervention might not occur until several years later.

8.2 Defining Comparison Groups

The evaluators all indicated that a key challenge was the lack of randomized design in the initiatives. As a result, they had to construct comparison groups for the initiative practices. Primary care practices and the health care system overall were undergoing a wide range of changes independent of the six primary care initiatives during the 2011–2017 time period when these initiatives were occurring. The fact that so many practices were changing their care delivery patterns, changing their organizational structure, and possibly participating in other

value-based purchasing initiatives (e.g., ACOs) made it difficult for evaluators to define comparison groups. Finding comparison regions was also a problem for some of the state initiatives like MAPCP and SIM. In these initiatives it was difficult to find comparison regions large enough to contain sufficient practices for analyses.

8.3 Data Challenges

Key informants were also challenged by data access problems. For example, evaluators for all the initiatives that included Medicaid patients reported significant delays in getting access to Medicaid data. As a result, Medicaid outcomes could not be included in most analyses. Another data challenge was the fact that practice consolidation during this period made it difficult to track individual practices over multiple years. Evaluators were also challenged by the number of practices that were joining shared savings programs sponsored by CMS or joining ACOs. Other data challenges for evaluators included tracking rolling entry into the initiatives and accounting for it in the evaluations and ensuring that patients were correctly assigned to the demonstration/comparison groups given that the tax identification numbers for many practices changed during the initiative.

9. Conclusions and Future Considerations

We focused on answering eight research questions related to the challenges that practices faced in transforming to advanced primary care practices, the strategies used to overcome these challenges, and the impact of the six initiatives. In Section 9.1 we summarize our findings, in Section 9.2 we discuss conclusions from our analyses, and in Section 9.3 we address study question 8 concerning implications for future CMMI testing and initiative scaling.

9.1 Summary of Key Findings

Practices made substantial progress transforming to advanced primary care practices. The evaluation reports, the analyses we conducted for this study using the ESF survey instrument developed for this project, and our interviews with key informants all indicate that initiative practices made large strides toward becoming PCMHs or advanced primary care practices. For example, while less than 10 percent of initiative FQHCs had any PCMH recognition status prior to the initiative, 70 percent of initiative FQHCs achieved NCQA Level-3 recognition by the end of the initiative. Similarly, the CPC evaluation found that CPC initiative practices improved their PCMH Assessment scores by about 50 percent during the CPC initiative. In an exploratory analysis we conducted for this meta-evaluation using data from the ESF and a simplified version of the SNMHS assessment, we found that, across the initiatives, evaluators' judgements about the practices they studied at the setting level indicated that practices had increased their practice transformation score by an average of more than 60 percent during the initiatives.

These initiatives produced modest and mixed results in the available evaluation period. Our review of evaluation results and our own analyses across the initiatives did not find consistent impacts across the initiatives or by setting within initiatives for any of the four core outcomes identified by CMS: fee-for-service Medicare hospital admissions, 30-day readmissions, outpatient ED visits, and Medicare expenditures. Of the 22 more granular initiative settings (seven CPC regions, FQHC as a whole, six HCIA-PCR awardees, and eight MAPCP states) for which we had cumulative results through Year 3, 10 settings experienced improvement relative to their comparison group for the cumulative three-year period for at least one of the four core outcome measures at a significance level of $p < 0.10$, and three of these settings (two CPC regions and HCIA TransforMED) experienced improvement on at least two core outcomes (see *Exhibit 5-3*).

As part of the systematic review, we also conducted pooled regression and meta-regression analyses across four initiatives (CPC, MAPCP, HCIA-PCR, and FQHC). The pooled regression and meta-regression analyses both indicated that the aggregate impacts of these four initiatives combined on the core outcomes were small and not statistically significant. For example, the aggregate change in Medicare's total cost of care was -\$26 per beneficiary per year (90% CI = -\$117 to \$66) in the meta-regression and -\$137 PBPY (90% CI = -\$310 to \$36) in the pooled regression. Although both results indicate a reduction in cost, neither was statistically significant. Similarly, the overall impacts on the other core outcomes (the three care utilization measures) were also not statistically significant and close to zero.

Across initiatives, certain population subgroups and practice types experienced more favorable outcomes. Using the pooled beneficiary-level data, we then attempted to

identify specific beneficiary- and practice-level characteristics that were associated with more favorable outcomes. We found that two beneficiary-level characteristics related to health status—disability status and HCC risk score quartile—both influenced the magnitude of the outcome effects, suggesting that the initiatives produced somewhat more favorable cost results for the sickest beneficiaries with the poorest health. Specifically, beneficiaries originally eligible for Medicare due to disability and beneficiaries with poor health (as reflected by being in the highest quartile of baseline HCC risk scores) experienced slower growth in Medicare expenditures. However, disability status and HCC risk score were not associated with statically significant impacts on overall rates of hospitalizations or ED visits, and non-dually eligible beneficiaries and those who were not originally eligible for Medicare due to disability experienced lower rates of 30-day readmissions.

Using the pooled beneficiary-level data, we also found slower growth in Medicare expenditures and lower rates of inpatient admissions and ED visits among practices with fewer than six practitioners and also among practices that were not multispecialty practices.

We then used meta-regression to test if different setting features were associated with better performance on each of the four core outcomes. From this analysis, we found that technical assistance provided by the initiatives had favorable effects on all outcomes except hospital admissions, although the relevant type of assistance differed by outcome. We conducted a formal Qualitative Comparative Analysis (QCA) for this study and identified two sets of initiative features that led to more favorable cost experience among 13 initiative settings that we could include in the analysis (seven CPC regions, FQHC, four MAPCP states, and HCIA Finger Lakes). These conditions were (1) having a high level of financial adequacy AND having a high level of staff adequacy AND not having a high level of practice participation in TA or (2) having leadership experience AND having a high level of practice usage of feedback reports AND having a high level of staff adequacy AND not having a high level of practice participation in TA. These results indicate that there were multiple paths to higher performance.

9.2 Conclusions

Based on our synthesis of the evaluation reports and the qualitative and quantitative analyses conducted for this study, we reached the following conclusions:

There are advantages to both state-convened and CMS-convened initiatives. Conveners are the entities responsible for organizing activities at the initiative level and can include state agencies, CMS, commercial payers, or health care systems. We found benefits and challenges to both state-convened and CMS-convened approaches (five of the six initiatives were convened by CMS or states). State agencies are able to capitalize on local relationships and understanding of local and state issues that might impact implementation and can adapt more easily to address those issues. However, state agencies also have limited resources and competing priorities that can make timely implementation of transformation activities challenging.

When feasible, a multi-payer design provides better support for practice transformation. A second key decision in convening a practice transformation initiative is whether it will be supported by a single payer or whether multiple payers will participate in the

initiative (multi-payer). Our study found that single-payer initiatives are certainly easier to administer for CMS and allow nationwide implementation of initiative designs and strategies that support practice transformation activities. However, practices typically want to transform for all their patients, so the level of financial support that a single payer typically can afford to provide is likely to be insufficient to pay for a practice's transformation activities. Multi-payer initiatives provide greater financial support and an opportunity to align transformation activities across more of a practice's patients. However, having multiple payers participate in an initiative does not guarantee they will align their views on the types of practice changes they desire, such as care coordination protocols and reporting of quality measures. As a result, multi-payer initiatives will be less burdensome on practices if the payers coordinate prior to implementation and collaborate to harmonize their requirements. CMS-convened multi-payer initiatives also limit CMS's ability to implement nationwide strategies because multi-payer efforts necessitate focusing on specific areas or regions that have other willing payers.

Regardless of convener or payer participation decisions, we also found that local exogenous factors and market characteristics can affect an initiative's implementation. Many ongoing reform efforts at the state and national level, as well as other changes in the complexity of health care markets, lead to additional stresses. Isolating a specific initiative's impact from the impacts of these confounding efforts and changes will continue to be a challenge for CMS as it designs, implements, and evaluates future demonstrations.

Practice-level factors are important in addressing transformation challenges. Our study found that practices often faced similar challenges in implementing transformation activities. We identified three major practice-level challenges to practice transformation: (1) a practice's inexperience with transformation activities, (2) limited health IT capabilities, and (3) difficulty integrating new care management staff into clinical activities. Practice capacity and readiness to implement practice transformation are important factors in practice-level change. Across all initiatives, evaluators noted that markers of practice capacity for change included past experience with health care transformation (most commonly achievement of NCQA PCMH accreditation), established relationships with organizations to deliver care coordination, and high-functioning health IT. While hiring new staff was a common transformation activity, key informants noted that practices needed to plan to integrate new staff into clinical workflow, ensure adequate time for implementation activities, and prepare for turnover. The evaluations found that advanced health IT capabilities were important, and health IT solutions needed to evolve to meet the demands of practice transformation, including supporting alternative payment models, care coordination, and patient-centered care. Two other findings related to practice transformation were that required benchmarks or milestones for the practice transformation process in CPC were valuable, but practices also said that a smaller number of milestones would be more feasible for them. Additionally, many practices were challenged by the uncertainty in identifying which of their patients were eligible for services and payments under an initiative, and practices want a process of attributing and identifying eligible patients that is prospective, transparent, and provides greater certainty.

Initiative-level supports also helped practices meet transformation challenges. Our study found that the level and type of financial support practices need may depend in part on their current level of readiness for, or progress toward, practice transformation. Particularly for small practices, up-front payments or higher payments at the beginning of the transformation

process may be necessary to support hiring care management staff, upgrading practice health IT capabilities, and changing treatment protocols. Shared savings incentives were problematic because their uncertainty and retrospective timing creates substantial financial risk for practices at the early stages of transformation. This is a particular problem for smaller practices that are less able or willing to bear risk. Financial support that combines shared savings with PBPM payments can overcome some of these challenges.

The six initiatives also used a variety of approaches for initiatives' learning systems, providing TA and other educational support to facilitate practice transformation. Our attempts to test the relationship between TA and outcomes were inconclusive. In our meta-regression analysis, some form of TA significantly contributed to impacts on three of the four outcomes, although the form varied by outcome. In contrast, our QCA analysis found that settings that performed better on Medicare cost savings actually had lower practice participation in TA activities (as judged by evaluators for each setting overall) than practices in lower-performing settings. This QCA finding may underscore the need for more tailoring of TA based on differences in practices' level of sophistication. More advanced practices sometimes viewed the available TA activities or resources as too basic to be useful to them, while the most tailored (and resource-intensive) form of TA—one-on-one direct assistance—often was offered predominantly to lower-performing practices that needed the most help. Thus, in designing an initiative's learning system, which is responsible for providing technical assistance, there is value in balancing the benefits of centralized oversight with the need to tailor TA on practice transformation opportunities for the circumstances of different regions or individual practices (such as tailoring TA for differences in practices' level of sophistication). Many key informants and evaluators reported that it is important for TA to start early, but mid-course adjustments are often necessary. Important topics for TA included health IT and the use of practice feedback reports.

Evaluators indicated that a majority of practices used the practice feedback data at least to a moderate extent to adjust their transformation efforts, but practices want feedback reports to contain actionable data. Practices often viewed claims-based data reports as untimely or as driven by factors beyond the influence of a primary care practice (e.g., specialist or hospital behaviors). The evaluators and key informants reported that practices thought feedback reports could be improved with aggregated, multi-payer data, more training on how data can be used and interpreted, and better incorporation of feedback reports into physician's normal workflow.

9.3 Implications for Scaling and Future Testing

This final section addresses implications for scaling and further testing. We categorized these implications into three areas:

9.3.1 Initiative Focus

- **Testing specific initiative features.** Our analyses indicate that some initiative features, such as TA, are associated with performance. However, the relationship between other features, such as the level of financial support or the type of feedback reports, and performance is less clear because this was not tested. Evaluations of future initiatives could provide CMS with a better understanding of the impact of certain features (such as TA and the level of financial support) on primary care

transformation. For example, CMS could provide different levels of financial support to one subset of practices and vary the type or amount of TA in a different subset.

- **Practice readiness.** In designing future initiatives, it may be valuable to test the impact of interventions focused on practices at a specific level of readiness. Broader initiatives could be tailored to different specific readiness levels within the initiative. Readiness in this context could include specific types of practice capacity, such as a practice's level of health IT or its level of resources to support transformation investments and activities (e.g., small practices versus large practices, or physician-owned practices versus practices owned by larger organizations). Alternatively, readiness could be defined based on whether the practice has already achieved formal PCMH recognition.
- **Small targeted initiatives.** There is also value in small, targeted initiatives, such as the HCIA PCR awards, that can provide smaller-scale test of concepts. However, in designing these initiatives, it is important to ensure they will have sufficient statistical power to detect meaningful impacts.
- **Operational focus.** In evaluating their initiatives, evaluators were often limited in assessing the marginal differences CMS funding made in improving primary care. While a valid research question, evaluations could be strengthened by also quantitatively assessing the contribution of different forms of operational guidance (e.g., TA, milestone requirements) and careful assessment of the extent of and type of transformation by practices.

9.3.2 Initiative Supports

- **Financial support.** Future testing could also vary specific types of financial support based on specific practice characteristics. For example, CMS could consider varying the level of PBPM support among smaller and larger practices, providing more up-front support for practices that are not yet PCMHs, or testing the impact of shared savings versus PBPM payments for larger practices only.
- **Harmonized requirements and measures.** Most key informants seemed to favor a multi-payer approach. Designing future multi-payer initiatives also provides an opportunity to align and harmonize requirements and measures among payers, aggregate data reports and data sources, and increase the influence of financial support mechanisms. CMS could consider the impact of requiring all payers to use consistent payment methods (as in MAPCP) versus allowing payers to use different methods (e.g., CPC).
- **Feedback data.** While many payers and practices want practices to be able to use feedback data to improve their care transformation activities, our review has highlighted multiple challenges in providing actionable feedback data. It would be valuable for follow-on initiatives to test whether these challenges, now recognized, can be sufficiently addressed with refinements to the existing framework of claims-based, payer-generated reports, or whether some alternative model of data provision to practices might be feasible.

- **Reducing administrative burden.** There is often a significant administrative burden for practices to participate in these initiatives, such as the documentation required to get NCQA PCMH recognition, and there could be value if future initiatives could test how a more streamlined set of administrative requirements that affect not only high-level outcome metrics (such as quality, cost and utilization, and patient satisfaction) but also more targeted metrics such as retention of participating practices and provider satisfaction.

9.3.3 Measures of Success

- **Initiative length.** For future initiatives that focus on practices that have not yet made much progress on transformation, a longer time period is likely necessary to measure the impact of their subsequent transformation activities under the initiative. For example, in FQHC half of the initiative practices that reached NCQA Level-3 PCMH recognition did not achieve recognition until the last quarter of the demonstration. More broadly, there was a consensus among key informants that the evaluations need to be conducted over a longer period of time, especially when transformation activities are expected to take multiple years.

While our meta-evaluation showed that primary care transformation was achieved by the majority of practices participating in the six initiatives, the impacts of these transformation activities on the core outcomes were limited in the relatively short period of time used for each initiative. Our study identified several key challenges to primary care transformation efforts that may limit or mask the resulting impacts. We believe these challenges can be addressed through careful planning and executing future large-scale and targeted evaluations.

References

- American Academy of Family Physicians [AAFP], American Academy of Pediatrics [AAP], American College of Physicians [ACP], and American Osteopathic Association [AOA]. (2007). Joint principles of the patient-centered medical home. Leawood, KS: American Academy of Family Physicians. Retrieved from http://www.aafp.org/dam/AAFP/documents/practice_management/pcmh/initiatives/PCMHJoint.pdf.
- Barr, M. S. (2010). The patient-centered medical home: Aligning payment to accelerate construction. *Medicare Care Research and Review*, 67(4):492-499. Retrieved from <http://journals.sagepub.com/doi/pdf/10.1177/1077558710366451>.
- Birnberg, J. M., Drum, M. L., Huang, E. S., Casalino, L. P., Lewis, S. E., Vable, A. M., ... Chin, M. H. (2011). Development of a Safety Net Medical Home Scale for Clinics. *Journal of General Internal Medicine*, 26(12), 1418–1425. <http://doi.org/10.1007/s11606-011-1767-9>
- Cubanski, J., & Neuman, T. (2016). The facts on Medicare spending and financing. Menlo Park, CA: The Henry J. Kaiser Family Foundation [KFF]. Retrieved from <http://files.kff.org/attachment/Issue-Brief-The-Facts-on-Medicare-Spending-and-Financing>.
- Edwards, S.T., Bitton, A., Hong, J., & Landon, B.E. (2014). Patient-centered medical home initiatives expanded in 2009-13: Providers, patients, and payment incentive increase. *Health Affairs*, 33(10), 1823-31.
- Gavin, N., Beil, H., Beadles, C., Witgert, K., Reck, J., Cousart, C., . . . Kissam, S. (2016). State Innovation Models [SIM] initiative evaluation: Model test year two annual report. Research Triangle Park, NC: RTI International.
- Gavin, N., Beil, H., Greenwald, L., West, S., West, N., Tant, E., . . . Hinkle, L. (2014). State Innovation Models [SIM] initiative evaluation: Model test base year annual report. Research Triangle Park, NC: RTI International.
- Ghosh, A., Kimmey, L., Li, E., Crosson, J., & Cheh, V. (2017). Draft interim evaluation report on the Independence at Home [IAH] demonstration. Princeton, NJ: Mathematica Policy Research.
- Gilman, B., Hoag, S., Moreno, L., Peterson, G., Barterian, L., Blue, L., . . . Zickafoose, J. (2014). Evaluation of Health Care Innovation Awards [HCIA] primary care redesign programs: First annual report. Princeton, NJ: Mathematica Policy Research.
- Glasgow, R. E., Vogt, T. M., & Boles, S. M. (1999). Evaluating the public health impact of health promotion interventions: The Reach Effectiveness Adoption Implementation Maintenance [RE-AIM] framework. *American Journal of Public Health*, 89(9). Retrieved from <https://www.ncbi.nlm.nih.gov/pmc/articles/PMC1508772/>.
- Hoff, T., Weller, W., & DePuccio, M. (2012). The patient-centered medical home: A review of recent research. *Medical Care Research and Review*, 69(6), 619-644.

Institute of Medicine. (2006). Constraints on sharing mental health and substance-use treatment information imposed by federal and state medical records privacy laws. In *Committee on crossing the quality chasm: Adaptation to mental health and addictive disorders* (Appendix B). Washington, DC: National Academies Press. Retrieved from <https://www.ncbi.nlm.nih.gov/books/NBK19829/>.

Jackson, G.L., Powers, B.J., Chatterjee, R., Bettger, J.P., Kemper, A.R., Hasselblad, V.,... Williams, J. W. (2013). The patient-centered medical home: A systematic review. *Annals of Internal Medicine*, 158(3), 169-78.

Kahn, K. L., Timbie, J. W., Friedberg, M. W., Mendel, P., Kress, A. M., Chen, E. K., . . . Setodji, C. M. (2016). Evaluation of CMS's Federally Qualified Health Center [FQHC] Advanced Primary Care Practice [APCP] demonstration: Final report evaluation. Santa Monica, CA: RAND Corporation.

Landon, B.E. (2014). Structuring payments to patient-centered medical homes. *The Journal of the American Medical Association*, 312(16), 1633-4.

Lipsey, M. W. (1990). *Design sensitivity: Statistical power for experimental research*. Newbury Park, CA: Sage Publications.

Magill, M. K., Ehrenberger, D., Scammon, D. L., Day, J., Allen, T., Reall, A. J., Sides, R. W., & Kim, J. (2015). The cost of sustaining a patient-centered medical home: Experience from 2 states. *The Annals of Family Medicine*, 13(5):429-435. Retrieved from <https://www.ncbi.nlm.nih.gov/pubmed/26371263>.

McCall, N., Haber, S., Van Hasselt, M., Cromwell, J., Sorensen, A., Farrell, K., . . . Yalowich, R. (2015a). Evaluation of the Multi-Payer Advanced Primary Care Practice [MAPCP] demonstration: First annual report. Research Triangle Park, NC: RTI International.

McCall, N., Haber, S., Van Hasselt, M., Cromwell, J., Sorensen, A., Farrell, K., . . . Yalowich, R. (2015b). Evaluation of the Multi-Payer Advanced Primary Care Practice [MAPCP] demonstration: Second annual report. Research Triangle Park, NC: RTI International.

Nichols, D., Haber, S., Romaine, M., Wiener, J. M., Leung, M., Smith, K., . . . Justice, D. (2017). Evaluation of the Multi-Payer Advanced Primary Care Practice [MAPCP] demonstration: Final report. Research Triangle Park, NC: RTI International.

Nichols, D., Haber, S., Romaine, M., Wiener, J. M., Leung, M., Smith, K., . . . Kinsler, S. (2016). Evaluation of the Multi-Payer Advanced Primary Care Practice [MAPCP] demonstration: Third annual report. Research Triangle Park, NC: RTI International.

Nielsen, M., Buelt, L., Patel, K., & Nichols, L. M. (2016). The patient-centered medical home's impact on cost and quality: Annual review of evidence 2014-2015. Washington, DC: Patient-Centered Primary Care Collaborative. Retrieved from <https://www.pcpcc.org/resource/patient-centered-medical-homes-impact-cost-and-quality-2014-2015>.

Nocon, R. S., Sharma, R., Birnberg, J. M., Ngo-Metzger, Q., Lee, S. M., & Chin, M. H. (2012). Association Between Patient-Centered Medical Home Rating and Operating Cost at Federally Funded Health Centers. *JAMA : The Journal of the American Medical Association*, 308(1), 60–66. <http://doi.org/10.1001/jama.2012.7048>

Peikes, D., Zutshi, A., Genevro, J. L., Parchman, M. L., & Meyers, D. S. (2012). Early evaluations of the medical home: Building on a promising start. *American Journal of Managed Care*, 18(2), 105-116.

Peikes, D., Anglin, G., Taylor, E. F., Dale, S., O'Malley, A., Ghosh, A., . . . Brown, R. (2016). Evaluation of the Comprehensive Primary Care [CPC] initiative: Third annual report. Princeton, NJ: Mathematica Policy Research.

Peikes, D., Ghosh, A., Zutshi, A., Taylor, E. F., Anglin, G., Converse, L., . . . O'Malley, A. (2016). Evaluation of the Comprehensive Primary Care [CPC] initiative: Second annual report. Princeton, NJ: Mathematica Policy Research.

Schneider, C. Q., & Wagemann, C. (2012). *Set-theoretic methods for the social sciences: a guide to qualitative comparative analysis*. Cambridge: Cambridge University Press.

Sinaiko, A.D., Landrum, M.B., Meyers, D.J., Alidina, S., Maeng, D.D., Friedberg, M.W., . . . Rosenthal, M.B. (2017). Synthesis Of research On patient-centered medical homes brings systematic differences into relief. *Health Affairs*, 36(3), 500-508.

Taylor, E. F., Anglin, G., Dale, S., Keith, R., Peikes, D., Shapiro, R., . . . Crosson, J. (2015). Evaluation of the Comprehensive Primary Care [CPC] initiative: First annual report. Princeton, NJ: Mathematica Policy Research.

The National Alliance for Health Information Technology. (2008). Report to the office of the national coordinator for health information technology defining key health information technology terms. Retrieved from <http://www.hitechanswers.net/wp-content/uploads/2013/05/NAHIT-Definitions2008.pdf>.

Wiley, J.A., Rittenhouse, D.R., Shortell, S.M., Casalino, L.P., Ramsay, P. P., Bibi, S., . . . Alexander, J. A. (2015). Managing chronic illness: Physician practices increased the use of care management and medical home processes. *Health Affairs*, 34(1), 78-86.