ABSTRACT In 2011 CareFirst BlueCross BlueShield, a large mid-Atlantic health insurance plan, implemented a payment and delivery system reform program. The model, called the Total Care and Cost Improvement Program, includes enhanced payments for primary care, significant financial incentives for primary care physicians to control spending, and care coordination tools to support progress toward the goal of higher-quality and lower-cost patient care. We conducted a mixed-methods evaluation of the initiative’s first three years. Our quantitative analyses used spending and utilization data for 2010–13 to compare enrollees who received care from participating physician groups to similar enrollees cared for by nonparticipating groups. Savings were small and fully shared with providers, which suggests no significant effect on total spending (including bonuses). Our qualitative analysis suggested that early in the program, many physicians were not fully engaged with the initiative and did not make full use of its tools. These findings imply that this and similar payment reforms may require greater time to realize significant savings than many stakeholders had expected. Patience may be necessary if payer-led reform is going to lead to system transformation.

Early Impact Of CareFirst’s Patient-Centered Medical Home With Strong Financial Incentives

Christopher C. Afendulis (afendulis@hcp.med.harvard.edu) is a senior research associate in the Department of Health Care Policy at Harvard Medical School, in Boston, Massachusetts.

Laura A. Hatfield is an assistant professor of health care policy and medicine in the Department of Health Care Policy, Harvard Medical School.

Bruce E. Landon is a professor of health care policy and medicine in the Department of Health Care Policy at Harvard Medical School and the Division of General Medicine and Primary Care at Beth Israel Deaconess Medical Center, both in Boston.

Jonathan Gruber is a professor of economics at the Massachusetts Institute of Technology, in Cambridge, Massachusetts.

Mary Beth Landon is a professor of health care policy in the Department of Health Care Policy, Harvard Medical School.

Robert E. Mechanic is a senior fellow at the Heller School for Social Policy and Management, Brandeis University, in Waltham, Massachusetts.

Darren E. Zinner is an associate professor at the Heller School for Social Policy and Management, Brandeis University.

By Christopher C. Afendulis, Laura A. Hatfield, Bruce E. Landon, Jonathan Gruber, Mary Beth Landrum, Robert E. Mechanic, Darren E. Zinner, and Michael E. Chernew

E fforts by insurers to transform the health care system have often focused on changing provider incentives. The broadest of these, population-based payment models, typically focus on primary care: Patients choose from a list of designated primary care providers or are assigned to a provider group based on their use of primary care services. Evaluations of population-based payment models are ongoing, but they generally have found positive, though often small, savings.\(^1\)\(^-\)\(^6\)

While alternative payment models create incentives for savings, realization of their potential requires transforming how care is delivered. It is not clear that delivery systems, particularly small primary care groups, are well suited to meet that challenge. Moreover, population-based payment models with downside risk (that is, financial penalties for failing to meet performance goals) might not be suitable for small practices (which tend to be more actuarially unstable than large practices). As a result, insurers have been experimenting with new models and making sizable investments to support delivery system reforms.

One such approach is the patient-centered medical home. This approach supports the delivery of whole-person, first-contact, coordinated care that emphasizes quality, safety, and enhanced access. As of 2013, more than 115 payer-sponsored patient-centered medical home demonstration programs throughout the United States were collectively caring for more than twenty million individuals.\(^7\) Early demonstrations (whose financial incentives for controlling costs were often weak or nonexistent) have shown mixed effects on total spending. While a
few demonstrations in integrated delivery systems or with shared-savings incentives have shown decreased utilization or spending, others across a range of settings have had little to no effect. In 2011 the largest health insurer in the mid-Atlantic region, CareFirst BlueCross BlueShield, launched the Total Care and Cost Improvement (TCCI) Program, an ongoing project that combines aspects of population-based payment and patient-centered medical home initiatives. The program includes enhanced payments for primary care, significant financial incentives for primary care physicians to control spending, and care coordination tools to support progress toward these goals. In this article we report the results from a mixed-methods evaluation of the program. Our quantitative evaluation focused on the first three years of this intervention. Our qualitative work explored how providers viewed the program and how it affected their behavior. The qualitative analysis provides context for our quantitative findings and aids in their interpretation.

Program Overview
The Total Care and Cost Improvement Program combines payment reform with delivery system support. On the financial side, CareFirst changed how it paid primary care physicians, giving them a 12 percent fee increase to compensate for extra care coordination and population management activities. CareFirst also set up a one-sided shared-savings system, in which primary care physician panels are given a budget target based on past spending. If panels spend less than their budget targets, they share in the savings. Instead of a lump sum distribution, shared savings are distributed via increased fees paid to the practice in the following year. For example, savings achieved in 2011 resulted in a twelve-month fee increase beginning July 1, 2012. These bonuses depend upon the savings realized by the panel, a quality score as assessed by CareFirst, panel size, and savings consistency over time. Beginning in July 2012, the average primary care fee increase (in addition to the 12 percent increase described above) was 25 percent; beginning in July 2013, it was 33 percent.

On the delivery system side, the Total Care and Cost Improvement Program hired more than seventy-five nurses in 2011 to coordinate care for the costliest patients—typically those with chronic illnesses. These care coordinators worked with physicians to develop care plans, and physicians were paid a fee ($170 in 2011) for each plan (these fees are included in our data). The program also provided physician panels with regular web-based reports of patient demographic characteristics, clinical patterns, spending, and risk profiles to identify candidates for care management.

Panels are required to have a minimum of 1,000 attributed patients and 5–15 primary care physicians. For the purposes of the program, CareFirst defined four panel types: “virtual panels,” which combined several solo or small practices; panels consisting of a single group; panels comprising portions of a single large physician group (which was divided to keep the number of physicians in the panel within the required range); and panels comprising portions of a panel based in a hospital or health system (which was also divided into smaller panels). For more information on panel types, see online Appendix A.

Patients enrolled in health maintenance organizations were required to choose a primary care physician and were attributed to that physician’s panel (although subsequent visits to another physician could change their panel attribution). Patients in preferred provider organization plans were assigned to a physician based on their use of primary care services during the previous twelve months. Any enrollee attributed to a primary care physician whose physician group participated in the Total Care and Cost Improvement Program was considered part of the program.

At its inception at the beginning of 2011, the Total Care and Cost Improvement Program had 2,014 primary care physicians in 653 practices, organized into 202 panels that collectively served 774,301 enrollees. By the end of 2013, the program consisted of 3,774 primary care physicians in 1,143 practices, organized into 282 panels that collectively served 1,181,548 enrollees.

Study Data And Methods

**Quantitative Analysis**

- **Data:** CareFirst provided enrollment data for the period 2010–13 that contained information on patients’ demographic characteristics, ZIP code of residence, risk score (calculated using the DxCG risk score model), insurance plan, and physician and panel assignments (for TCCI participants only) by month. The claims data for this period included spending, place of service, physician identifiers, diagnosis codes, and Current Procedural Terminology codes.

- **Study Sample and Variables:** Our unit of observation was the patient-quarter. We required that each enrollee be enrolled with CareFirst from the first quarter of 2010 through the first quarter of 2011. We excluded enrollees who were younger than nineteen or older than sixty-four,
There was not an overall sense that the program compelled a majority of the physicians to focus on cost containment.

(For a description of this alternative regression approach, see Appendix C.)

As noted above, we analyzed five dependent variables: total spending, inpatient spending, outpatient spending, primary care physician evaluation and management visits, and specialist evaluation and management visits. Supplemental analyses tested for differential effects by panel type.

All models were run using ordinary least squares regression, and standard errors were clustered at the panel level. All quantitative analyses were performed with Stata, version 13.0.

**Sensitivity analyses:** We conducted five sensitivity analyses. First, we expanded the treatment group to include TCCI enrollees who started later in 2011 than the first quarter and whose panels started in the program in the same quarter as they did. This more inclusive definition of program participation captures much of the crossover from participants to nonparticipants during our study period. Second, we assessed the impact of excluding enrollees who were not continuously enrolled with CareFirst from 2010 to 2013. Third, we considered the impact of excluding enrollees who crossed over from participants to nonparticipants (or vice versa). Fourth, we estimated a two-part model that explicitly modeled the probability of any spending in a given quarter and log spending, using subjects with any spending.

Finally, following the approach of Alison Cuellar and coauthors, we modified our two-part model. We replaced enrollee fixed effects with enrollee traits, expanded our sample in a number of ways, defined participation in the Total Care and Cost Improvement Program as having ever been in the program, allowed the effect of being in the program to grow over time, and modified the functional form of the regression models. (For details about this alternative regression approach, see Appendix D.)

**Qualitative analysis** We conducted thirty-
It is difficult for global payment models to yield total savings, especially in markets with many small practices.

nine semistructured interviews across twenty primary care panels between January and May 2014. We interviewed thirty primary care physicians, four individual practice administrators, and five executives from larger panels. Of the physicians interviewed, twelve were in virtual panels, ten were in single-group panels, and eight were in panels that were part of larger medical groups. (Further details about the panel selection are available in Appendix E.)

We conducted sixty-minute semistructured interviews with informants that focused on their general understanding of the Total Care and Cost Improvement Program, perceptions of the value of specific program components (such as nurse care coordinators and performance reports), perceptions of changes in their practices in response to the program, and perceived barriers to performance improvement.

During each interview, a team member assisting the interviewer took detailed notes, which were reviewed by the interviewer immediately after the interview was completed. The notes were entered into a qualitative software program (Atlas.ti) and analyzed for thematic content. Members of the project team identified initial themes for the queries and developed additional themes based on a review of the queries.

LIMITATIONS There were a number of limitations to our analysis. First, our intent-to-treat strategy included only primary care physicians and their patients who joined the Total Care and Cost Improvement Program at its onset. This mitigates concern about nonrandom assignment to the program, but it left out many physicians and patients. It also created crossover, which could have reduced our observed savings. Our sensitivity analyses that included additional panels that started in the program in 2011 addressed this concern.

Second, nonrandom physician participation may be a concern. Third, our results show an average effect. Some practices may have responded more strongly than others did to the program.

Fourth, the generalizability of our results is limited by the geographic concentration of the CareFirst population in Maryland, the District of Columbia, and Virginia. Maryland in particular has a number of unique policy features (all-payer rate setting, for example).

Finally, the qualitative analysis included relatively few subjects. Nevertheless, if one believes that panels that were less engaged with the program would be less likely to participate in our survey, then our qualitative findings of relatively modest effects, despite this selection, would likely remain.

Study Results

QUANTITATIVE FINDINGS We observed only minor differences between the Total Care and Cost Improvement Program and control groups at baseline in terms of risk score, number of chronic conditions, age, sex, relationship to primary enrollee, and enrollment in risk-based contracts (Exhibit 1). The two groups also displayed only small differences in total per quarter baseline spending ($1,014 for the TCCI group versus $1,038 for the control group). Program enrollees had slightly more primary care physician visits and slightly fewer specialist visits, relative to nonenrollees.

We also assessed the degree of crossover between participant and nonparticipant status among the enrollees. More than two-thirds of those in the TCCI group (215,132 out of 298,463) stayed in the program for the duration of their enrollment in CareFirst (data not shown). Among the nonenrollees, more than half (283,133 out of 537,778) remained outside of the program for the remainder of their CareFirst tenure. Of those who crossed over into the program, more than one-third (72,525 out of 182,120) did so during its “ramp-up” period in 2011, in the same quarter as their program panel began operating. They were therefore included in our first sensitivity analysis sample.

SPENDING AND UTILIZATION ANALYSES: Unadjusted raw total spending patterns were very similar for enrollees and nonenrollees in the Total Care and Cost Improvement Program in the periods both before and after the program’s implementation (Exhibit 2). Average quarterly spending for both groups increased from about $1,000 to $1,300 over the four-year period, but there was no observable change in spending between the two groups after the initiation of the program.

Adjusted results from our regression models confirmed these findings. We found no signifi-
cant differences in overall spending (which included bonus payments and program costs) associated with the Total Care and Cost Improvement Program between enrollees and nonenrollees (adjusted difference between groups per quarter: $2.24 per quarter; 95% confidence interval: −$14.31 to $18.78) (Exhibit 3). (Full results for all regression specifications and dependent variables are available in Appendix H.)

For all years except 2010, these spending figures included the Total Care and Cost Improvement Program–based fee increases provided to primary care physicians. According to information from CareFirst, these fee increases accounted for 1.1 percent of total spending for program enrollees during 2011–13, reaching 3 percent of spending in 2013.

The relative change in inpatient spending was positive ($5.28) and not significant (Exhibit 3). The relative change in outpatient spending was negative (a reduction of $3.04) but small relative to the average outpatient spending amount of $794 and also not significant. Our measures of utilization, while admittedly coarse, also found little impact of the program. There was a reduction in primary care visits (0.0079 visits compared to a mean of 0.44 visits), but the change was not significant. There was also a very small increase in specialist visits (0.01 visits compared to a mean of 0.75 visits), but that too was not significant.

▸ SENSITIVITY ANALYSES: As noted above, we conducted a variety of sensitivity analyses and found their results to be substantively similar to those of our base analysis. (Appendix Exhibit G.1 summarizes the results for all the sensitivity analyses, with the exception of the alternative two-part model, whose results are presented in Appendix Exhibit G.2. Full regression results are in Appendix H.)

The results from the second regression model specification, which included numerous interaction terms for the impact of the Total Care and Cost Improvement Program over time, are summarized in Appendix F. We plotted regression-adjusted trends in each of the five outcome variables. We observed few differences in outcomes between program enrollees and nonenrollees, either before or after the program’s implementation (Appendix Exhibits F.1 and F.2). We can reject the joint null hypothesis that all of the interaction terms were 0 for outpatient spending (F-test statistic: 4.05; \(p < 0.001\)), primary care physician visits (F-test statistic: 19.35; \(p = 0.001\)), and specialist visits (F-test statistic: 26.47; \(p < 0.001\)). However, in all three cases, there was neither a consistent pattern of lower spending or utilization nor any substantively significant difference between the two groups.

EXHIBIT 1

Selected characteristics of CareFirst BlueCross BlueShield’s Total Care and Cost Improvement Program enrollees and nonenrollees, first quarter 2011

<table>
<thead>
<tr>
<th></th>
<th>Enrollees (n = 298,463)</th>
<th>Nonenrollees (n = 537,778)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Risk score</td>
<td>1.09</td>
<td>1.10</td>
</tr>
<tr>
<td>One chronic condition</td>
<td>0.94%</td>
<td>0.95%</td>
</tr>
<tr>
<td>Two or more chronic conditions</td>
<td>6.4%</td>
<td>6.5%</td>
</tr>
<tr>
<td>Age (years)</td>
<td>45.0</td>
<td>45.1</td>
</tr>
<tr>
<td>Female</td>
<td>59%</td>
<td>59%</td>
</tr>
<tr>
<td>Primary enrollee</td>
<td>67%</td>
<td>68%</td>
</tr>
<tr>
<td>Risk contract</td>
<td>49%</td>
<td>40%</td>
</tr>
<tr>
<td>Total spending per quarter</td>
<td>$1,014</td>
<td>$1,038</td>
</tr>
<tr>
<td>Inpatient spending per quarter</td>
<td>$317</td>
<td>$333</td>
</tr>
<tr>
<td>Outpatient spending per quarter</td>
<td>$697</td>
<td>$705</td>
</tr>
<tr>
<td>Visits to primary care provider per quarter</td>
<td>0.45</td>
<td>0.37</td>
</tr>
<tr>
<td>Visits to specialists per quarter</td>
<td>0.61</td>
<td>0.64</td>
</tr>
</tbody>
</table>

SOURCE Authors’ analysis of CareFirst BlueCross BlueShield claims and enrollment data for 2010–11.

NOTES Risk score based on DxCG risk score model. Total spending excludes prescription drug spending.

EXHIBIT 2

Total spending for Total Care and Cost Improvement Program enrollees and nonenrollees, by quarter, 2010–13

SOURCE Authors’ analysis of CareFirst BlueCross BlueShield claims and enrollment data for the period 2010–13.

NOTE Total spending excludes prescription drug spending.
authors. They used a broader sample (including beneficiaries who joined CareFirst after 2011), a different definition of patient-centered medical home participation, and a different specification and found greater savings. When we approximated their sample and model, we found comparable savings. This result seems driven largely by the inclusion of beneficiaries joining CareFirst after 2011 and the replacement of fixed effects with enrollee traits, although other modeling decisions (such as how missing risk scores were treated and propensity score weighting) might play a role as well (for a description of this alternative regression approach, see Appendix D).17

**Qualitative Findings** Our qualitative findings generally were consistent with the quantitative ones. While the thirty primary care physicians we interviewed generally expressed support for the goals of the program, there was scant evidence of substantial cost-reducing behavior change. Most of the respondents indicated that they had not changed how they practiced because of the Total Care and Cost Improvement Program, and many of those who had changed were focused on quality more than on cost savings. For example, physicians were reluctant to make major changes to their referral patterns but were more receptive to making quality improvements.

Several factors suggest why substantial cost-saving behavior change was hard. First, most physicians were overwhelmingly focused on their clinical activities and had limited exposure to administrative or financial matters. As a result, they did not consistently seek out or use information or tools provided by CareFirst. For example, very few physicians used the initial version of the online reporting portal, describing it as difficult to use. Most reviewed reports profiling their high-risk patients only when the reports were delivered to their offices by nurse care coordinators or presented at quarterly meetings by CareFirst staff.

Second, many physicians did not understand, and in some cases claimed that they did not care about, the program’s largest financial opportunity: the shared-savings incentive. Only about 20 percent of the physicians we interviewed could accurately describe how the savings award was calculated, and only 25 percent said the potential to earn a bonus offered a strong enough motivation to practice differently. About a quarter of the respondents said that they did not keep track of spending because finances were the responsibility of others in the practice. Several respondents said that they had received an award but were unsure what their panel or they individually had done to warrant the award.

This is not to say that the respondents did not view the program favorably. The primary care physicians were generally positive about the nurse care coordinators, and physicians in small practices particularly valued the care coordination support. They felt that the nurses provided concrete benefits to their patients and provided valuable information about patients between office visits. But there was not an overall sense that the program compelled a majority of the physicians to focus on cost containment.

**Discussion**

Our results suggest that the CareFirst Total Care and Cost Improvement Program was largely budget-neutral during its first three years, when fees and bonuses are included. It is important to recognize that in contrast to evaluations of other payment models, where bonus payments were made separately from normal claims payments, the CareFirst bonuses were implemented as percentage increases in the fee schedule. Therefore, our spending measure included bonus pay-
ments. CareFirst reported that bonuses were approximately 1 percent initially, growing to about 3 percent. This result (that savings were offset by bonuses) was similar to the findings from early evaluations of other population-based payment models—such as the Blue Cross Blue Shield of Massachusetts Alternative Quality Contract, which did not generate total savings until year four.5

We believe that over the long run, pre-bonus spending is more important than total spending because it reflects actual service use. Lower service use will allow slower increases in benchmarks over time. While it is tempting for payers to try to capture savings very quickly, there will be no incentive for practices to transform themselves if savings are not shared. While the sharing of savings (which should be expected in a shared-savings program) will weaken the financial profile of the program from a payer or societal perspective, impatience on the part of payers may hinder providers’ improvement.

In the case of the Total Care and Cost Improvement Program, the effects on evaluation and management services were small and generally not significant, so the small pre-bonus savings suggested by our results would have to reflect changes in use of other services. Our qualitative findings suggest that during the early period of the program, many providers neither fully understood it nor cared about the financial incentives. Moreover, the mechanism for rewarding efficient care (increasing payment rates) may have had less salience than an immediate lump sum payment. Finally, the care plans and case managers were largely focused on improving the health of their patients instead of reducing inefficient or low-value care.

The number of care plans in the program was quite low during our study period (peaking at about 4,000 in December 2013, which represented approximately one plan for each physician participating in the program). Hiring and care plan activity increased in 2014–15, which may have affected utilization. The program may have had important benefits for enrollees with care plans, but the limited number of cases in our data makes it difficult to find any effect of the program. If the program improved quality, it did so without increasing spending.

CareFirst has acknowledged early growing pains with the program—in particular, in terms of physician education and the recruitment and retention of nurse care coordinators. Those early challenges led to fewer care plans being written than CareFirst had projected, and lower levels of physician engagement than anticipated. Since then, CareFirst has reduced the turnover rate among care coordinators and has ramped up education efforts aimed at physicians in the panels. As providers better understand program incentives and information, particularly related to referrals, the impact of the Total Care and Cost Improvement Program may grow. Providers are facing a variety of payment reform initiatives, so there may also be opportunities for payers to coordinate efforts to improve providers’ understanding of and response to payment incentives. It may be the case that multipayer initiatives will focus provider attention more effectively than efforts by a single payer can.

Conclusion
This analysis illustrates that it is difficult for global payment models to yield total savings, especially in markets with many small practices. It may be that the effectiveness of the Total Care and Cost Improvement Program will increase over time. Modifications to the program could also increase its impact, including making bonuses more salient and focusing more on efficient care (for example, by encouraging referrals to lower-cost specialists).

Ultimately, payers must find ways to improve quality at a sustainable cost. Although the Total Care and Cost Improvement Program has many attractive features, further experience along with experimentation and modification are needed to achieve the goal of significantly higher-quality, lower-cost care. Clearly, observers who anticipated immediate benefits from payment reform were overly optimistic. Though early results show some savings (pre-bonus), patience may be needed until the full potential of payment reform is better understood.

---

CareFirst BlueCross BlueShield provided funding for this project. Jonathan Gruber acknowledges a long-standing consulting arrangement with CareFirst. Michael Chernew serves on an advisory board of the National Institute for Health Care Management and is an adviser to Archway Health. He holds equity in VIBID Health.
NOTES


7 Edwards ST, Bitton A, Hong J, Landon BE. Patient-centered medical home initiatives expanded in 2009–13: providers, patients, and payment incentives increased.


10 Maeng DD, Khan N, Tomcavage J, Graf TR, Davis DE, Steele GD. Reduced acute inpatient care was largest savings component of Geisinger Health System’s patient-centered medical home. Health Aff (Millwood). 2015;34(4):636–44.


15 CareFirst. Program description and guidelines for CareFirst Patient-Centered Medical Home Program (PCMH) and Total Care and Cost Improvement Program (TCCI) [Internet]. Baltimore (MD): CareFirst; 2016 Oct 1 [cited 2017 Jan 3]. Available from: https://provider.carefirst.com/carefirst-resources/provider/pdf/pcmh-program-description-guidelines.pdf

16 CareFirst evaluated each panel’s quality in five domains: physician engagement with the program, appropriate use of services (a subset of the preventable quality indicators developed by the Agency for Healthcare Research and Quality), effectiveness of care (the Healthcare Effectiveness Data and Information Set developed by the National Committee for Quality Assurance), patient access, and structural capabilities.

17 To access the Appendix, click on the Appendix link in the box to the right of the article online.

18 Because some enrollees had a prescription drug benefit carved out of their CareFirst insurance contract, we did not have drug claims for a subset of enrollees. For this reason, we excluded drug spending from the analysis.