

Working Paper Series  
Congressional Budget Office  
Washington, D.C.

Lessons from Medicare's Demonstration Projects  
on Value-Based Payment

Lyle Nelson  
Health and Human Resources Division  
Congressional Budget Office  
([Lyle.Nelson@cbo.gov](mailto:Lyle.Nelson@cbo.gov))

January 2012  
Working Paper 2012-02

To enhance the transparency of CBO's work and to encourage external review of it, CBO's working paper series includes papers that provide technical descriptions of official CBO analyses and papers that represent original, independent research by CBO analysts. Working papers are not subject to CBO's regular review and editing process. Papers in this series are available at [www.cbo.gov/publications](http://www.cbo.gov/publications) (select Working Papers).

I am grateful to Melinda Buntin, James Baumgardner, Philip Ellis, Linda Bilheimer, Holly Harvey, Tom Bradley, Jean Hearne, Randall Brown, Michael Chernew, David Cutler, and Mary Kapp for helpful comments and suggestions. Michael Levine and Sarah Axeen assisted with the tables.

## Lessons from Medicare's Demonstration Projects on Value-Based Payment

### **Abstract**

This paper summarizes the results of Medicare demonstrations of value-based payment systems, which give providers financial incentives to improve the quality and efficiency of care. Only one of the four demonstrations for which results are available has yielded significant savings for the Medicare program. In that demonstration, Medicare made bundled payments to hospitals and physicians to cover all services connected with heart bypass surgeries, and Medicare spending for those services declined by about 10 percent. The other demonstrations appear to have resulted in little or no savings for Medicare. One, the Physician Group Practice Demonstration, allowed large multispecialty physician groups to share in estimated savings if they reduced total Medicare expenditures for their patients. Another offered hospitals bonuses if they met certain criteria regarding the quality of care. The last (for which results are available only on a preliminary basis for the first year) allowed home health agencies to share in estimated savings if they reduced total Medicare expenditures for their patients and met certain targets for quality of care.

## Introduction

Driven by an aging population and rising health care costs, federal spending on Medicare is projected to exert increasing strains on the federal budget over the coming decades. At the same time, there are widespread concerns about the quality of health care and the efficiency with which care is delivered to beneficiaries. Those concerns are not unique to Medicare but rather extend to the U.S. health care system more generally. They have prompted widespread interest in policies that many experts believe could improve the quality of care for patients while reducing costs. Prominent among such policies are value-based payment systems, which alter the financial incentives for health care providers in ways that are designed to improve the quality of care or the efficiency of its delivery.

This paper reviews the key findings of demonstrations that have tested two broad approaches to value-based payment in Medicare's fee-for-service program:

- Pay-for-performance, in which an insurer's payments to providers depend in part on whether the providers meet certain targets regarding the quality or efficiency of care, and
- Bundled payment, in which an insurer makes a single, comprehensive payment that covers services furnished by multiple providers during a defined episode of care.

In general, pay-for-performance seeks to improve care by financially rewarding providers who deliver high-quality or efficient care, and bundled payments seek to align the incentives of different providers in a way that is designed to yield greater efficiency and coordination of care. Both approaches have been adopted or tested by private health plans.

Evaluation results are available for four Medicare demonstrations of value-based payment, and other demonstrations are under way or planned. The evaluations show that only one of the four value-based payment demonstrations has resulted in significant savings for Medicare.<sup>1</sup> That demonstration, in which Medicare made bundled payments to cover all hospital and physician services for heart bypass surgeries, reduced Medicare spending on those services by about 10 percent (see Table 1). The other demonstrations appear to have resulted in little or no savings for Medicare. One demonstration, the Physician Group Practice (PGP) Demonstration, offered large multispecialty physician groups a share in the savings if they reduced total Medicare expenditures for their patients; another demonstration offered hospitals bonuses if they met certain criteria regarding the quality of care. The latter demonstration ended, but the PGP demonstration continues with a somewhat modified design that has incorporated some key lessons from the original demonstration.<sup>2</sup> Preliminary results for a fourth demonstration, which

---

<sup>1</sup> The findings of another set of Medicare demonstrations, which tested disease management and care coordination programs, are discussed in detail in Lyle Nelson, [Lessons from Medicare's Demonstration Projects on Disease Management and Care Coordination](#), Working Paper 2012-01 (January 2012). For additional discussion of both sets of demonstrations, see Congressional Budget Office, [Lessons from Medicare's Demonstration Projects on Disease Management, Care Coordination, and Value-Based Payment](#), Issue Brief (January 2012).

<sup>2</sup> The modified design for the PGP demonstration became effective in January 2011. The results discussed in this paper pertain to the original demonstration.

**Table 1.**

<b>Key Features and Results of the Demonstrations of Value-Based Payment</b>				
	<b>Pay-for-Performance</b>			<b>Bundled Payment</b>
	<b>Physician Group Practice Demonstration</b>	<b>Premier Hospital Quality Incentive Demonstration</b>	<b>Medicare Home Health Pay-for-Performance Demonstration</b>	<b>Medicare Participating Heart Bypass Center Demonstration</b>
Estimated Effect on Medicare Expenditures	Little or no effect on expenditures	No effect on expenditures	Little or no effect on expenditures in the first year	10 percent reduction in expenditures on bypass surgeries
Estimated Effect on Quality of Care	Small improvement in processes of care	Small improvement in processes of care	Little or no effect on patient outcomes in the first year	Little or no effect on patient outcomes
Payment Approach	Group practices received bonuses if they reduced total Medicare spending on their patients. The bonuses depended partly on the number of quality of care targets met.	Hospitals received bonuses if they met certain quality of care targets.	For each region, estimated savings were distributed to home health agencies that had the highest quality scores or the greatest improvement in quality scores	Hospitals and physicians received bundled payments for heart bypass surgeries
Participating Organizations	10 physician group practices	278 hospitals	556 home health agencies in 7 states	7 hospitals and the physicians on their medical staffs who treated heart bypass patients
Randomized Design?	No	No	Yes	No
Period of Operation	2005–2010	2003–2009	2008–2009	1991–1996

Source: Author's review of information in the sources listed in Appendix A.

offered home health agencies a share in the savings if they reduced total Medicare expenditures for their patients and met certain targets for quality, do not show a reduction in Medicare spending in the first year. Other value-based payment demonstrations currently under way for which results are not yet available are discussed in Appendix B.

The Patient Protection and Affordable Care Act (PPACA, Public Law 111-148) included numerous provisions that are intended to improve the quality and efficiency of health care services delivered under Medicare, some of which build on approaches that were tested in the demonstrations discussed in this paper. One such provision established a shared savings program that is similar in many respects to the approach tested in both the original and modified PGP demonstration. Under that program, physicians and other providers that voluntarily form “accountable care organizations” will receive bonuses if the Medicare spending on their patients is below a certain target and if they meet quality-of-care requirements. Other provisions of

PPACA that build on approaches tested in the demonstrations include programs in which hospitals and certain other providers will be eligible to receive additional payments based on the quality of care they deliver, and a national pilot program in which Medicare will make bundled payments that cover services received by beneficiaries during an episode of care surrounding a hospitalization. PPACA also established the Center for Medicare and Medicaid Innovation within the Centers for Medicare & Medicaid Services (CMS); that provision gives the Secretary of Health and Human Services (HHS) broad authority to develop and test new approaches to paying health care providers and delivering health care benefits. The Secretary has the authority to implement such approaches on a broader scale (including nationally) if they are found to have favorable effects on quality of care and expenditures. Although the programs that will be implemented broadly under PPACA and those that will be tested by the innovation center (and might later be implemented broadly) could yield favorable results, the findings from earlier demonstrations suggest that many challenges may arise in implementing such programs in ways that reduce Medicare spending. The demonstrations hold important lessons for the design and implementation of those programs.

## Pay for Performance

In recent years, a growing number of private health plans have implemented pay-for-performance programs that use financial incentives to encourage physicians and hospitals to provide high-quality care. At the same time, Medicare's administrators have undertaken a variety of demonstrations to test those approaches and assess their impact on enrollees' care and federal spending. Results from independent evaluations are available for two major demonstrations that tested those approaches:

- In the Physician Group Practice Demonstration, 10 large physician group practices could share some of the savings if they reduced total Medicare spending on their patients. The bonuses received by the physician groups depended in part on the proportion of quality-of-care targets they met.
- In the Premier Hospital Quality Incentive Demonstration, 278 hospitals were eligible to receive bonuses if their quality scores exceeded certain thresholds.

Those two demonstrations are discussed separately in this paper because they are very different from each other. A shorter discussion is provided for a third demonstration—the Medicare Home Health Pay-for-Performance Demonstration—because findings from the evaluation are available for only the first year.

### Pay-for-Performance Programs in the Private Sector

The design of pay-for-performance programs in the private sector varies greatly in terms of how quality of care is measured and how the financial incentives are structured. One or more of four types of quality measures are typically included in pay-for-performance programs. Measures of the *process of care* capture whether a patient's care is consistent with practice guidelines (for example, whether diabetic patients receive annual foot and eye exams). Measures of the *outcomes of care* include intermediate clinical outcomes (such as the results of blood pressure and cholesterol tests) as well as health outcomes (such as mortality or complications after surgery). *Structural measures* capture the types of structures and policies in place (such as the

adoption of health information technology). Measures of *patient satisfaction* capture patients' perceptions on issues such as the ability of providers to communicate clearly and show respect. Many pay-for-performance programs also include measures of the cost of care, in some cases narrowly defined, such as the extent to which physicians prescribe generic drugs when they are available.

Pay-for-performance programs vary in how they determine whether a provider receives a financial reward or bonus. The approaches fall into three categories: rewarding providers whose quality scores exceed a specified threshold, rewarding those whose scores are in the top tier of eligible providers (in the top 10 percent, perhaps), and rewarding providers based on the amount of improvement in their scores. Each approach offers advantages and disadvantages. Rewarding improvement creates stronger incentives for many providers—particularly those whose quality scores are initially very low—to improve. But providers who already have very high quality scores might have limited room for improvement and are likely to regard a system in which bonuses are based solely on improvement as unfair. Recognizing those trade-offs, some programs use a combination of approaches in which bonuses are made partly on the basis of improvement and partly on reaching a certain benchmark for quality.

Primary care physicians are the most common type of provider included in pay-for-performance programs, although specialists and hospitals have been included to a growing extent in recent years.<sup>3</sup> Such programs are more common among health maintenance organizations (HMOs) than among preferred provider organizations. HMOs are more likely to use pay for performance in contracts with physicians if they pay physicians (or medical groups) by capitation and if they require that enrollees select a primary care physician who is responsible for authorizing all referrals.<sup>4</sup>

### **Physician Group Practice Demonstration**

The Medicare, Medicaid, and SCHIP Benefits Improvement and Protection Act of 2000 (P.L. 106-554) mandated a demonstration to test the effects of providing financial incentives to physician group practices to improve the quality and efficiency of care delivered to their Medicare patients. Ten PGPs were selected for the demonstration, ranging in size from 230 to 1,300 physicians (see Table 2). Two were freestanding group practices, seven were part of an integrated delivery system that included at least one hospital, and one was a network consisting of 60 small practices.<sup>5</sup> The groups that were in two of the integrated delivery systems were faculty group practices within an academic medical center. The demonstration began in April 2005 and originally was scheduled to last three years. It was subsequently extended for another two years, ending in March 2010. All 10 PGPs subsequently chose to participate in the Physician

---

<sup>3</sup> Meredith B. Rosenthal and others, "Climbing Up the Pay-for-Performance Learning Curve: Where Are the Early Adopters Now?" *Health Affairs*, vol. 26, no. 6 (November/December 2007), pp. 1674–1682.

<sup>4</sup> Meredith B. Rosenthal and others, "Pay for Performance in Commercial HMOs," *New England Journal of Medicine*, vol. 355, no. 18 (November 2, 2006), pp. 1895–1902.

<sup>5</sup> The physician network is supported by a management services organization that provides quality improvement, information management, contracting, and other services to independent physician practices, each of which was offered the choice to participate in the demonstration.

**Table 2.**

<b>Participants in the Physician Group Practice Demonstration</b>			
<b>Participant</b>	<b>Geographic Location</b>	<b>Type of Service Area</b>	<b>Number of Physicians</b>
Billings Clinic	Montana/Wyoming	Small city, rural	232
Dartmouth-Hitchcock Clinic	New Hampshire/Vermont	Small city, rural	907
Everett Clinic	Washington	Suburban, small city	250
Forsyth Medical Group	North Carolina	Small city	250
Geisinger Clinic	Pennsylvania	Small city, rural	833
Marshfield Clinic	Wisconsin	Small city, rural	1,039
Middlesex Health System	Connecticut	Suburban, small city	293
Park Nicollet Clinic	Minnesota	Large metropolitan, suburban	648
St. John's Clinic	Missouri/Arkansas	Small city, rural	522
University of Michigan Faculty Group Practice	Michigan	Suburban, small city	1,291

Source: Department of Health and Human Services, *Report to Congress: Physician Group Practice Demonstration Evaluation Report* (2009).

Group Practice Transition Demonstration, which began in January 2011 under a design that was modified in some respects to incorporate lessons learned from the original demonstration. The PGPs will be given the option of joining the Medicare Shared Savings Program that is scheduled to be implemented in 2012 or a shared savings initiative to be implemented by the innovation center.

**Design.** Under the demonstration, the participating PGPs were paid on a fee-for-service basis and were eligible to receive a bonus if the total Medicare expenditures for their patients were at least 2 percent below a target.<sup>6</sup> The target for each PGP was intended to represent the Medicare expenditures that would have been incurred for its patients in the absence of the demonstration, and the 2 percent threshold was specified to avoid paying bonuses for small differences from the target that may have been the result of random variation in Medicare expenditures. The PGPs were accountable for the *total* Medicare expenditures of their patients, including expenditures for services provided by other physicians, inpatient hospital care, and postacute care.

---

<sup>6</sup> Information on the design of the demonstration is from Department of Health and Human Services, *Report to Congress: Physician Group Practice Demonstration First Evaluation Report* (2006), [www.cms.gov/reports/downloads/Leavitt1.pdf](http://www.cms.gov/reports/downloads/Leavitt1.pdf); and John Kautter and others, “Medicare Physician Group Practice Demonstration Design: Quality and Efficiency Pay-for-Performance,” *Health Care Financing Review*, vol. 29, no. 1 (Fall 2007), pp. 15–29. Appendix A lists all of the studies of the demonstrations I reviewed for this paper.

PGPs could receive up to 80 percent of the estimated savings in Medicare expenditures for their patients, defined as the difference between target expenditures and actual expenditures beyond the 2 percent threshold. The Medicare program retained the rest of the savings. PGPs received the entire 80 percent of the estimated savings if they met all of a specified set of quality-of-care targets. Otherwise, they received a fixed portion of the 80 percent and an additional amount that depended on the proportion of the quality targets met. (The share of the bonus pool that PGPs received regardless of their quality-of-care scores declined from 70 percent in the first year of the demonstration to 50 percent in the final year.) PGPs did not face a penalty if the expenditures for their patients were above the target. However, if the expenditures for a PGP's patients in a given year were more than 2 percent higher than the target, the amount of those higher expenditures beyond the 2 percent threshold was used to offset estimated savings for that PGP in subsequent years.

Beneficiaries did not enroll in the PGPs and continued to have the same freedom of provider choice that they had before the demonstration. The set of patients for whom each PGP was held accountable in a given year was determined at the end of the year through an analysis of Medicare claims. In particular, beneficiaries were retroactively assigned to a PGP if they received more evaluation and management services from the PGP than from any other physician practice.<sup>7</sup> Beneficiaries were not eligible to be assigned to a PGP if they were receiving hospice services at the beginning of the year or if they were enrolled in a Medicare Advantage plan at any time during the year. The roughly 220,000 beneficiaries who were assigned to the participating PGPs each year received an average of 85 percent of their evaluation and management services from those PGPs. The number of assigned beneficiaries varied across PGPs from roughly 10,000 to 40,000.

The expenditure target for each PGP in a given year was determined retroactively by updating the average Medicare expenditures for the PGP's assigned beneficiaries during a base year before the demonstration to reflect the estimated growth in expenditures that would have occurred for those beneficiaries absent the demonstration. The growth rate in that calculation was the growth in expenditures during the period for a comparison group of beneficiaries who lived in the PGP's service area but who received no evaluation and management services from the PGP. Beneficiaries in the comparison group were required to have received at least one evaluation and management service from another provider. In addition, beneficiaries were not included in the comparison group for a given year if they had been assigned to the PGP in a previous year, if they were receiving hospice services at the beginning of the year, or if they were enrolled in a Medicare Advantage plan during the year. The expenditures for the PGP patients and the comparison group were risk-adjusted using the hierarchical condition categories (HCC) model to remove the effects of any changes in health status in the two populations.<sup>8</sup>

---

<sup>7</sup> The assignment of beneficiaries to PGPs was based on the allowed charges for evaluation and management services delivered in an office or other outpatient setting. Evaluation and management includes the services provided during physician visits, such as examination of patients and counseling, but it does not include surgical procedures or tests.

<sup>8</sup> CMS uses the HCC model to adjust payments to Medicare Advantage plans to account for differences in enrollees' health status. The HCC model uses information on beneficiaries' diagnoses and demographic characteristics to predict Medicare expenditures.

Under that approach, the determination of whether the expenditures for a PGP's patients were above or below the target in a given year depended on whether the growth in expenditures for its patients from the base period was above or below the growth for the comparison group (after removing the effects of any differences between the two groups in the change in risk scores). The expenditures for a PGP's patients were below the target if they grew more slowly than did those of the comparison group and above the target if the opposite was true. The calculation did not depend on how the *level* of spending for a PGP's patients compared with that of its comparison group. Thus, although a PGP may have been a lower-cost provider than others in the area before the demonstration, such a difference did not directly affect whether the PGP was above or below its targets during the demonstration.

A total of 32 quality-of-care measures were specified for the demonstration: 10 for diabetes, 10 for congestive heart failure (CHF), 7 for coronary artery disease (CAD), 3 for hypertension, and 2 for preventive care. Only the diabetes measures were used in the first year of the demonstration, the CHF and CAD measures were added in the second year, and the others were added in the third year. Twenty-five of the measures were constructed from data in the medical records of PGP patients, and 7 were constructed from Medicare claims. For each measure, a PGP could meet the target by satisfying one or more of three criteria: achieving at least 75 percent compliance, achieving certain scores relative to national scores for Medicare Advantage plans, or closing at least 10 percent of the gap between the PGP's value before the demonstration and 100 percent compliance.

**Implementation.** All of the participating PGPs developed or expanded disease management or care coordination programs for the demonstration. In some cases, PGPs implemented programs for the demonstration that had been developed previously for other populations. The programs primarily targeted patients with CHF and diabetes as well as certain high-risk patients, typically defined as those with multiple chronic conditions or heavy service use. Nurses employed by the PGPs served as care managers, and the main focus of the programs was patient education and monitoring. Three PGPs provided their CHF patients with home electronic monitoring devices, several PGPs enhanced their follow-up and monitoring of patients who had been discharged from a hospital, and several developed and promoted palliative care programs for terminally ill patients.<sup>9</sup> The costs of implementing such programs were borne by the PGPs, which had to cover those costs out of their revenues (including their bonuses from Medicare).

All of the PGPs had experience with care management prior to the demonstration, either as a result of operating an HMO or because they had been delegated such functions under contracts with health plans. The PGPs developed some new care management programs for the demonstration and also applied programs they had developed for other patient populations to their Medicare fee-for-service patients.

---

<sup>9</sup> Information on the implementation of the demonstration is from Department of Health and Human Services, *Report to Congress: Physician Group Practice Demonstration First Evaluation Report* (2006); and Michael Trisolini and others, *Medicare Physician Group Practices: Innovations in Quality and Efficiency* (paper prepared by RTI International for the Commonwealth Fund, December 2006), [www.commonwealthfund.org/usr\\_doc/971\\_Trisolini\\_Medicare\\_physician\\_group\\_practices\\_i.pdf](http://www.commonwealthfund.org/usr_doc/971_Trisolini_Medicare_physician_group_practices_i.pdf).

Most of the PGPs had electronic medical record systems before the demonstration, and those that did not were in the process of developing them. None of the PGPs undertook major information technology initiatives as a result of the demonstration but rather adapted and expanded their systems. For example, most developed patient registries that identified their Medicare patients who had specific chronic conditions and included key clinical information for those patients, such as the results of laboratory tests, which were used by the care managers.

The PGPs reported that they tried to improve the quality of care through various other means, such as encouraging physicians to follow evidence-based practice guidelines and providing feedback to physicians on how their performance compares with that of their peers. Those activities were generally under way at the PGPs before the demonstration, however, and the extent to which they were intensified during the demonstration is not clear.

**Key Findings.** Two PGPs received bonuses in the first year of the demonstration because Medicare expenditures for their patients were more than 2 percent below their target, four PGPs received bonuses in the second year, five received bonuses in the third and fourth years, and four received bonuses in the fifth year.<sup>10</sup> Two PGPs received bonuses in all five years. Those results, although informative, do not provide a measure of the effects of the demonstrations on Medicare expenditures. More detailed information on the effects of the demonstration is contained in a Report to the Congress by the Secretary of Health and Human Services, which covers the first two years of the demonstration, but even that report presented a limited set of results.<sup>11</sup> That analysis concluded that, during the second year of the demonstration, total Medicare expenditures for the patients of the 10 PGPs were about 1 percent less than the targets (corresponding estimates were not presented for individual PGPs or for other years).<sup>12</sup> After accounting for the bonuses paid to the PGPs, the estimated net savings for the Medicare program for the second year of the demonstration was about \$1.6 million.<sup>13</sup> That represents a net savings of about \$7 per beneficiary across the 10 participating PGPs—a net reduction in Medicare spending of about 0.1 percent. Estimated savings were lower for the first year.

There are two reasons why that estimate might overstate the amount of any net Medicare program savings. First, it appears that some PGPs changed their diagnostic coding practices under the demonstration in a way that increased the risk scores of their patients relative to those of the comparison group. Such a change would increase the risk-adjusted expenditure targets, making it easier for PGPs to appear to achieve savings (since the targets are adjusted to account for changes from the base period in the relative risk scores of the PGPs' patients and the

---

<sup>10</sup> Centers for Medicare & Medicaid Services, *Medicare Physician Group Practice Demonstration: Physician Groups Continue to Improve Quality and Generate Savings Under Medicare Physician Pay-for-Performance Demonstration* (July 2011) [www.cms.gov/DemoProjectsEvalRpts/downloads/PGP\\_Fact\\_Sheet.pdf](http://www.cms.gov/DemoProjectsEvalRpts/downloads/PGP_Fact_Sheet.pdf).

<sup>11</sup> Department of Health and Human Services, *Report to Congress: Physician Group Practice Demonstration Evaluation Report* (2009), [www.cms.gov/reports/downloads/RTC\\_Sebelius\\_09\\_2009.pdf](http://www.cms.gov/reports/downloads/RTC_Sebelius_09_2009.pdf). The estimates presented in the report are based on analyses conducted by Research Triangle Institute.

<sup>12</sup> Ibid., p. 9.

<sup>13</sup> HHS estimated the net federal savings as the government's share of the savings beyond the 2 percent threshold for the PGPs that received bonuses minus the amount of any increase in expenditures at other PGPs beyond the 2 percent threshold. PGPs that were within 2 percent of their target were assumed to have had no effect on Medicare expenditures.

comparison group). From the base year to the second year of the demonstration, the risk scores increased by an average of 8 percent for the PGPs' patients, compared with an average increase of 5 percent for the comparison group.<sup>14</sup> The risk scores for PGP patients increased faster than those for the comparison group in 9 of the 10 PGPs, including all four that received bonus payments. The evaluation did not estimate how much of the greater increase in risk scores for PGP patients was the result of changes in coding practices. However, several PGPs reported that they began encouraging their physicians to code all appropriate diagnoses on claims under the demonstration to help identify and manage the care of patients with chronic conditions and to improve the accuracy of their risk-adjusted expenditure targets.<sup>15</sup> That indicates that at least some of the greater increase in risk scores for PGP patients was attributable to changes in coding practices. Such a result is not unexpected when payments to providers are made in part on the basis of the patients' risk scores.

A second factor that might have led to an overestimate of Medicare program savings is that, for the four PGPs that received bonuses in the second year of the demonstration, Medicare expenditures per patient were growing more slowly than were those of the local comparison groups before the demonstration.<sup>16</sup> This is important because, in determining expenditure targets, the growth in expenditures for the comparison groups was used as a proxy for the growth that would have occurred for the PGPs' patients in the absence of the demonstration. When the targets for the demonstration's second year were adjusted to account for the differences in preexisting expenditure trends, Medicare expenditures at the four PGPs that received bonuses were not significantly different from the targets. It is not known what caused the lower spending trends for the four PGPs before the demonstration or whether they would have continued in the absence of the demonstration. To the extent that those lower spending trends would have continued, the savings estimated for those PGPs are overstated. For the six PGPs that did not receive bonuses, spending trends before the demonstration were similar for PGP patients and the comparison groups.

The PGPs met an average of nearly 90 percent of the quality-of-care targets in the first year of the demonstration, when only the 10 diabetes measures were in effect. In the second year, when 10 CHF measures and seven CAD measures were added, the PGPs met an average of 98 percent of the targets. Most quality measures were constructed from data in the medical records of PGP patients, and the evaluation did not collect corresponding data for the comparison groups. Consequently, the only quality measures that could be compared for PGP patients and the comparison groups are the seven constructed from claims data. For four of those measures, PGP patients had somewhat greater improvement in measured quality of care from the base year to the second year of the demonstration, with the estimates implying that the demonstration increased the proportion of PGP patients who received certain recommended tests by amounts

---

<sup>14</sup> There has been a general increase in HCC risk scores in the Medicare fee-for-service program that has been attributed primarily to increased accuracy and specificity in the diagnostic coding on claims.

<sup>15</sup> Department of Health and Human Services, *Report to Congress: Physician Group Practice Demonstration First Evaluation Report* (2006), pp. 58–59.

<sup>16</sup> Department of Health and Human Services, *Report to Congress: Physician Group Practice Demonstration Evaluation Report* (2009), pp. 56–62.

ranging from 1 to 5 percentage points. PGP patients and the comparison groups had similar changes in quality for the other three measures.

### Premier Hospital Quality Incentive Demonstration

CMS initiated a pay-for-performance demonstration for hospitals in response to an unsolicited proposal from Premier, a company owned by approximately 200 nonprofit hospitals and health systems. Premier provides data management, purchasing, and other services to hospitals and other health care providers. The objective of the demonstration was to test the effects of providing hospitals with financial incentives to provide high-quality care. Eligibility for the demonstration was limited to the 444 hospitals that were reporting quality measures to other payers through Premier's quality measurement system as of March 2003; 278 hospitals chose to participate. The demonstration began in October 2003 and was originally scheduled to last three years. CMS subsequently extended the project for an additional three years, and the demonstration ended in September 2009.

**Design.** Under the demonstration, hospitals reported data on the quality of care provided to patients in five clinical areas: acute myocardial infarction (AMI), CHF, pneumonia, coronary artery bypass graft surgery, and hip or knee replacement. That information was used to compute a composite quality score for each hospital for each clinical area. During each of the first three years of the demonstration, hospitals with quality scores in the highest 10 percent of the distribution for a given clinical area received a bonus equal to 2 percent of their Medicare payments for patients in that clinical area, and hospitals with quality scores in the second-highest 10 percent received a 1 percent bonus. In addition, hospitals with quality scores in the top 50 percent for any of the 5 clinical areas received special recognition on the CMS web site. Hospitals were not subject to penalties in the first two years. In the third year, hospitals with quality scores at or below the threshold corresponding to the bottom 20 percent of the distribution of first-year scores were required to pay a penalty of 2 percent (for those in the lowest 10 percent) or 1 percent (for those in the second-lowest 10 percent). Thus, hospitals that initially had low scores could avoid penalties if they improved their scores sufficiently. Hospitals also could avoid a penalty by dropping out of the demonstration. By the end of the third year, nearly 10 percent of the hospitals that began the demonstration had withdrawn.

The incentive structure was changed for the second three years of the demonstration to provide bonuses based on a broader set of criteria. Beginning in the fourth year, 40 percent of the amount that CMS budgeted for bonus payments was awarded to hospitals whose composite quality scores were at or above the median of the distribution of scores from two years earlier. Hospitals were eligible for additional bonuses if their quality scores for the year were in the top 20 percent or if their improvement in scores from two years earlier was in the top 20 percent. Hospitals were required to pay a penalty if their quality scores for the year were below the threshold corresponding to the bottom 20 percent of scores from two years earlier.

In all, 34 quality measures were defined for the five clinical areas. Twenty-seven were process measures that reflect the percentage of patients whose treatment followed recommended practices and 7 were outcome measures that reflect the percentage of patients who did not experience certain adverse outcomes (such as death or complications from surgery). The number of quality measures per clinical area varied from 4 to 9.

The demonstration did not use a randomized design. The evaluation estimated the effects of the demonstration on quality of care and Medicare payments by comparing outcomes for patients at demonstration hospitals with outcomes for patients at other hospitals, using statistical methods to control for differences in the characteristics of the hospitals that are expected to affect outcomes.

Other initiatives to improve the quality of hospital care were under way during the demonstration, and average reported quality scores were improving nationwide. The Hospital Quality Alliance (sponsored by organizations representing the hospital industry, insurers, employers, and consumers) began encouraging hospitals to voluntarily report data on their quality of care in 2003. The percentage of hospitals that reported such data increased dramatically in 2004 as a result of a requirement established in the Medicare Prescription Drug, Improvement, and Modernization Act of 2003 (P.L. 108-173) that hospitals report certain quality of care measures each year in order to receive the full update in their Medicare payment rates. The hospital quality scores are made available each year to beneficiaries on an HHS Web site ([www.hospitalcompare.hhs.gov](http://www.hospitalcompare.hhs.gov)). Twenty measures of quality were specified for the first year of the national reporting requirement; 18 were used in the demonstration. In addition, some private insurers had pay-for-performance programs in effect for hospitals during the demonstration. All of those efforts probably contributed to a nationwide trend of rising quality scores.

**Key Findings.** The composite quality scores for the demonstration hospitals rose from the first quarter of the demonstration to the twelfth quarter for each of the five clinical areas.<sup>17</sup> On average, the increase ranged from 7 percentage points for AMI (the average score rose from 90 percent to 97 percent) to 22 percentage points for CHF (the average score rose from 67 percent to 89 percent). Those increases in composite quality scores were primarily the result of improvements in the process scores. There was little room for improvement in the outcome scores, which were very high at the start of the demonstration.

Not all of the increases in quality scores for the demonstration hospitals can be attributed to the effects of the demonstration, because quality scores were improving nationwide. The best available evidence indicates that the demonstration was responsible for small increases in quality of care and that most of the increases in quality that occurred at the participating hospitals would have occurred in the absence of the demonstration. The CMS-funded evaluation reported that the effect of the demonstration was to raise quality scores during the three-year period by an average of about 1 to 4 percentage points.<sup>18</sup> Another study that used somewhat different methods

---

<sup>17</sup> Centers for Medicare & Medicaid Services, *Evaluation of the Premier Hospital Quality Incentive Demonstration. Executive Summary: Impacts on Quality of Care, Medicare Reimbursements, and Medicare Beneficiaries' Length of Stay During the First Three Years of the Demonstration* (March 2009), [www.cms.gov/reports/downloads/Premier\\_ExecSum\\_2010.pdf](http://www.cms.gov/reports/downloads/Premier_ExecSum_2010.pdf). Estimates in the report were based on analyses conducted by Abt Associates, under contract to CMS.

<sup>18</sup> Ibid., pp. 12–15.

obtained similar estimates.<sup>19</sup> There is little information available on the effects of the demonstration on health outcomes. One study reported that the demonstration had no effect on beneficiary mortality rates within 30 days of hospital admission, but no information is available for other health outcomes.<sup>20</sup>

During the first three years of the demonstration, CMS paid bonuses to hospitals that totaled nearly \$25 million. The total amount of the bonuses was about 0.25 percent of the Medicare payments to the demonstration hospitals for admissions in the five clinical areas during the three-year period. The penalties in the third year were about \$0.1 million. Few hospitals were required to pay a penalty because their quality scores had increased sufficiently by the third year. That increase in quality scores was generally consistent with a nationwide trend during the demonstration period. In addition, hospitals could avoid paying a penalty by withdrawing from the demonstration.

The demonstration did not affect Medicare expenditures for inpatient hospital care. In principle, if the demonstration improved the quality of care, it could have reduced Medicare expenditures for inpatient care by reducing the number of readmissions. Another possible mechanism is a reduction in the prevalence of complications that result in patients being assigned to a higher-cost diagnosis-related group (DRG). The CMS-funded evaluation examined this issue by comparing Medicare expenditures for “episodes of care” initiated at demonstration hospitals with those initiated at other hospitals, using statistical methods to control for differences in the characteristics of the hospitals, their market areas, and their Medicare patients.<sup>21</sup> (The evaluation defined an episode as beginning with an admission for one of the five clinical areas and extending to 90 days after discharge. Episodes include readmissions to any hospital during that period.) The analysis showed that the demonstration did not affect Medicare expenditures for inpatient hospital care during the defined episodes of care and, after accounting for the bonuses that were paid to some hospitals, there was no discernible net effect on Medicare expenditures. (The point estimates in that analysis were a 0.1 percent increase in expenditures on inpatient care and a 0.3 percent net increase in expenditures, after accounting for the bonuses. Neither estimate was significantly different from zero at the 5 percent level.) The analysis also showed that the demonstration did not affect the total number of days that patients were hospitalized during the episodes of care or the number of days they were hospitalized for readmissions. Another study that used similar methods also concluded that the demonstration had no effect on Medicare expenditures.<sup>22</sup> In principle, the demonstration could have affected Medicare expenditures over a

<sup>19</sup> Peter K. Lindenauer and others, “Public Reporting and Pay for Performance in Hospital Quality Improvement,” *The New England Journal of Medicine*, vol. 356, no. 5 (February 1, 2007), pp. 486–496. Both the CMS-funded evaluation and the study reported by Lindenauer and colleagues estimated the effects of the demonstration on quality of care by comparing quality scores for the demonstration hospitals with quality scores that other hospitals submitted to CMS for Hospital Compare. The analysis was limited to three clinical areas (AMI, CHF, and pneumonia) because those were the areas for which data were available from Hospital Compare.

<sup>20</sup> Andrew M. Ryan, “Effects of the Premier Hospital Quality Incentive Demonstration on Medicare Patient Mortality and Cost,” *Health Services Research*, vol. 44, no. 3 (June 2009), pp. 821–842.

<sup>21</sup> Centers for Medicare & Medicaid Services, *Evaluation of the Premier Hospital Quality Incentive Demonstration. Executive Summary*.

<sup>22</sup> Andrew M. Ryan, “Effects of the Premier Hospital Quality Incentive Demonstration on Medicare Patient Mortality and Cost.”

longer period of time and for other services, such as physicians' services and post-acute care. However, given the evidence that the demonstration had limited effects on quality of care, noticeable effects on Medicare expenditures over a longer time period or for other services are unlikely.

### **Medicare Home Health Pay-for-Performance Demonstration**

The Medicare Home Health Pay-for-Performance Demonstration was initiated by CMS to test the effects of providing financial incentives to home health agencies (HHAs) to improve quality of care. Under the demonstration, HHAs received their standard payments from Medicare and were eligible to receive additional payments if the participating HHAs in their region achieved Medicare savings and if they met certain criteria regarding the quality of care. The demonstration began in January 2008 and lasted two years. Findings from the evaluation are currently available for the first year of the demonstration.

**Design.** A total of 556 HHAs in seven states that volunteered to participate were randomly assigned either to a treatment group or to a control group. For each of four regions, CMS estimated the Medicare savings attributable to the demonstration and distributed the entire amount of those savings to HHAs in the treatment group whose quality scores were in the top 20 percent among demonstration participants in their state or whose improvement in quality scores was in the top 20 percent.<sup>23</sup> Thus, although it was anticipated that the demonstration might lead to improvements in the quality of care that could reduce Medicare expenditures, the demonstration was not intended to yield net savings for the Medicare program. HHAs in regions that did not achieve savings did not receive any additional payments, regardless of their quality scores.

To estimate the effect of the demonstration on Medicare expenditures in each region, CMS compared the actual expenditures for patients of the HHAs in the treatment group with an estimate of what those expenditures would have been without the demonstration. Those latter amounts were estimated by updating the average expenditures for patients in the treatment HHAs during a base year before the demonstration by the growth in average expenditures over that period for patients of the control group HHAs. The analysis included expenditures on all Part A and Part B Medicare services incurred by beneficiaries during each home health episode and up to 30 days after the episode.

If actual expenditures for patients of the treatment HHAs in a given region were lower than projected expenditures by any amount, the difference was considered as savings and was returned to the treatment HHAs in that region. In contrast to the PGP demonstration, the estimated savings did not need to exceed a threshold to be considered as savings. If “negative savings” were estimated for one or more regions, those amounts were not used to offset the savings in other regions before determining bonus payments to HHAs.

---

<sup>23</sup> Participants were located in seven states divided into four regions: Northeast (Connecticut and Massachusetts), South (Alabama, Georgia, and Tennessee), Midwest (Illinois), and West (California).

Seven quality measures (a subset of the 21 measures that every HHA must submit under Medicare) were used to determine the HHAs' eligibility for bonuses: the number of hospital admissions among the HHA's patients; the use of emergency care; and improvements in bathing, ambulation, transferring from one setting to another, management of oral medications, and the status of surgical wounds. Bonus payments were determined separately for each measure. The demonstration did not require any additional data collection or submission by the participating HHAs.

**Key Findings.** During the first year, the demonstration had no overall effect on the seven measures of the quality of care that were used to determine HHAs' eligibility for bonuses and had little or no effect on the other reportable measures.<sup>24</sup> Those results measured the total effect of the demonstration for all of the HHAs that participated. The results for individual states were more varied but nevertheless indicate that the demonstration had no overall effect on quality of care. Of the 147 state-level estimates (reflecting 21 measures for each of 7 states), about two-thirds indicated that the demonstration had no effect, and the remaining estimates were nearly equally divided between those that indicated an improvement in quality and those that indicated a reduction in quality.

An analysis of the effects of the demonstration on Medicare expenditures showed that HHAs in three regions achieved combined savings of about \$15 million during the first year; expenditures were about \$9 million higher than projected in the fourth region.<sup>25</sup> However, the estimated savings were very small as a proportion of projected spending—ranging from 0.4 percent in one region to 1.2 percent in another region—suggesting that those differences between actual and projected expenditures could reflect random variation in Medicare expenditures rather than actual savings. The absence of a notable effect on Medicare expenditures is consistent with the finding that the demonstration had no discernible effect on quality of care—particularly in the number of hospitalizations and in the use of emergency care.

## Bundled Payments

Bundled-payment arrangements are generally viewed as arrangements in which a single payment from an insurer covers services that typically are furnished by two or more providers. Such arrangements are not commonly used by private insurers, primarily because of the challenges in designing and implementing bundled-payment systems.<sup>26</sup> Results are available from a single

<sup>24</sup> David F. Hittle, Eugene J. Nuccio, and Angela A. Richard, *Evaluation of the Medicare Home Health Pay-for-Performance Demonstration. CY2008 Report—Volume 1: Agency Characteristics, Costs, and Quality Measure Performance among Treatment, Control, and Non-Participant Groups* (report prepared by the Division of Health Care Policy and Research, University of Colorado Anschutz Medical Campus for the Centers for Medicare & Medicaid Services, June 2011), [www.cms.gov/reports/downloads/HHP4P\\_Demo\\_Eval\\_2008\\_Vol1.pdf](http://www.cms.gov/reports/downloads/HHP4P_Demo_Eval_2008_Vol1.pdf).

<sup>25</sup> The effects of the demonstration on Medicare expenditures were not estimated as part of the evaluation but instead were estimated by the research firm that assisted CMS in designing the demonstration. See Alan White and Henry Goldberg, *Home Health Pay-for-Performance: Medicare Savings Calculations* (memorandum submitted by Abt Associates to the Centers for Medicare & Medicaid Services, January 26, 2010). That analysis estimated effects on total Part A and Part B expenditures during each home health episode and for up to 30 days afterward.

<sup>26</sup> See Mark McClellan, “Reforming Payments to Healthcare Providers: The Key to Slowing Healthcare Cost Growth While Improving Quality?” *Journal of Economic Perspectives*, vol. 25, no. 2 (Spring 2011), pp. 69–92.

independent evaluation of one major demonstration in Medicare that tested bundled payment: For the Medicare Participating Heart Bypass Center Demonstration, Medicare made bundled payments that covered all inpatient hospital and physician services for coronary artery bypass graft surgeries performed at seven hospitals.

### Bundled-Payment Programs in the Private Sector

The bundled-payment arrangements in the private sector most similar to those in the Medicare demonstration examined in this paper involve organ transplants—although such arrangements have been adopted to a limited extent by private insurers for other services. A recent study by the Government Accountability Office (GAO) found that all five of the largest national health insurers—Aetna, Cigna, Humana, UnitedHealth Group, and WellPoint—have been making bundled payments for organ transplants for more than 20 years.<sup>27</sup> Each of those insurers has developed “transplant networks” consisting of hospitals that meet certain criteria regarding the number of transplants performed and the quality of care. The insurers make bundled payments to those hospitals that cover all hospital and physician services for all phases of the transplant episode—namely, the evaluation, the transplant procedure, and follow-up care (including readmissions) for periods ranging from 30 to 365 days. All of the insurers attempt to steer patients to the hospitals in their transplant networks by reducing patients’ cost sharing if they use those hospitals. One insurer does not cover transplant services unless they are provided at a hospital that is part of its transplant network.

More broadly, bundled payment means paying providers a single payment for a set of services rather than paying separately for each service. By that definition, private and public insurers use a number of bundled payment approaches in their fee-for-service payment systems. For example, private insurers typically pay per diem rates for inpatient hospital care, and Medicare pays a set amount for each stay that depends on the patient’s diagnosis and whether certain major procedures are performed. Moreover, Medicare and most private insurers pay physicians global fees for surgeries that cover the surgery and pre- and postoperative care over specified periods.

Two of the five insurers in the GAO study make bundled payments for services other than organ transplants. One makes such payments for bariatric surgery at selected hospitals in 22 states, and another makes bundled payments for coronary artery bypass graft surgery (also called heart bypass surgery) and other selected cardiovascular procedures at a single hospital. In both cases, the bundled payments cover all hospital and physician services required for the surgery and follow-up care over a specified period. Four of the five insurers are considering developing bundled-payment arrangements for other services, such as cardiac and orthopedic surgery.

The five major insurers in the GAO study have chosen to make bundled payments for organ transplants because they are expensive procedures used by a small proportion of their enrollees, and transplant episodes have clearly defined starting and ending points. The insurers reported that several challenges have hindered them from implementing bundled payments on a broader scale. All process claims for bundled services manually because modifying their automated

---

<sup>27</sup> Government Accountability Office, *Medicare: Private Sector Initiatives to Bundle Hospital and Physician Payments for an Episode of Care*, GAO-11-126R (January 2011), [www.gao.gov/new.items/d11126r.pdf](http://www.gao.gov/new.items/d11126r.pdf).

claims-processing systems to handle bundled payments would be expensive. In addition, the insurers reported that they have encountered difficulty negotiating bundled-payment contracts with providers, although some providers have recently become more receptive to such arrangements. Finally, insurers noted that defining episodes of care for most conditions is challenging.

Two other private initiatives are the bundled-payment systems for heart bypass surgery and certain other procedures developed by the Geisinger Health System in Pennsylvania and the Prometheus Payment Model, an approach to bundled payment developed for certain procedures and for the management of certain chronic conditions.<sup>28</sup> The Prometheus model was developed by a nonprofit organization (the Health Care Incentives Improvement Institute) primarily with support from the Robert Wood Johnson Foundation. The bundled payments are based on the resources that are required for treating a particular condition, based on clinical guidelines and expert opinion, and are adjusted to account for differences in patient severity.<sup>29</sup> The payments to providers also depend in part on their scores on quality-of-care measures. The payment model is being tested at four sites.

### **Medicare Participating Heart Bypass Center Demonstration**

CMS initiated the Medicare Participating Heart Bypass Center Demonstration to test the effects of bundling hospital and physician payments for coronary artery bypass graft surgery. Four hospitals began the demonstration in 1991, and three others were added in 1993. Participation of the four original hospitals was scheduled for three years, and the period was extended for an additional two years. The three later-starting hospitals operated for three years. CMS selected the 7 hospitals from 27 applicant institutions.<sup>30</sup> CMS's criteria for selecting participants for the demonstration included the projected Medicare savings implied by their proposed bundled-payment rates, the quality of care at the hospital, and the volume of bypass surgeries performed there.

**Design.** Under the demonstration, Medicare made bundled payments for bypass surgery that covered all services provided by the hospital, the surgeon, and other physicians.<sup>31</sup> For the surgeon, the bundled payments covered preoperative care, the surgery, and postoperative care in the hospital and at the surgeon's office, just as global surgical fees cover such services in the

---

<sup>28</sup> For a discussion of the bundled-payment systems developed by Geisinger, see Douglas McCarthy and others, *Geisinger Health System: Achieving the Potential of System Integration Through Innovation, Leadership, Measurement, and Incentives* (Commonwealth Fund, June 2009), [www.commonwealthfund.org/Publications/Case-Studies/2009/Jun/Geisinger-Health-System-Achieving-the-Potential-of-System-Integration.aspx](http://www.commonwealthfund.org/Publications/Case-Studies/2009/Jun/Geisinger-Health-System-Achieving-the-Potential-of-System-Integration.aspx).

<sup>29</sup> Health Care Incentives Improvement Institute, *PROMETHEUS Payment: On the Frontlines of Health Care Payment Reform* (Robert Wood Johnson Foundation, July 2010), [www.rwjf.org/files/research/66748.pdf](http://www.rwjf.org/files/research/66748.pdf).

<sup>30</sup> Bypass surgery was performed at more than 700 hospitals in the U.S. when CMS (at the time known as the Health Care Financing Administration) issued the solicitation for the demonstration. In all, 206 institutions submitted preapplications expressing interest. CMS invited 42 of that group to submit formal applications, and 27 did so.

<sup>31</sup> Information on the design of the demonstration is from Jerry Cromwell and others, *Medicare Participating Heart Bypass Center Demonstration: Volume I, Final Report* (report prepared by Health Economics Research for the Health Care Financing Administration, July 1998), [www.cms.gov/demoprojectsevalrpts/MD/ItemDetail.asp?ItemID=CMS063472](http://www.cms.gov/demoprojectsevalrpts/MD/ItemDetail.asp?ItemID=CMS063472).

Medicare fee-for-service program. For other physicians, the bundled payment covered only the services provided to bypass patients during their hospitalization. The bundled payments did not depend on the patient's severity of illness or on the volume and intensity of services provided. The hospitals agreed to forgo the outlier payments they would have received for extraordinarily costly cases under Medicare's inpatient prospective payment system; instead, an allowance for expected outlier payments was incorporated into the bundled payment. The bundled payment also covered any readmissions for related conditions for periods that varied by hospital from three days to six weeks after discharge.

The demonstration was intended to align the incentives of hospitals and physicians and lead to greater efficiency and coordination among providers. Under the Medicare fee-for-service program, hospitals and physicians are paid separately and face different incentives. Hospitals receive a prospectively determined payment for each stay that depends on a patient's diagnosis and, in some cases, on whether certain procedures were performed. In contrast, physicians generally receive a separate payment for each service they provide. Moreover, unlike hospitals, physicians do not face a direct financial incentive to economize on hospital-provided services because they do not bear the costs of such services.

The bundled-payment rates for the demonstration were determined through negotiations between the applicants and CMS. Applicants were required to propose separate rates for two DRGs in which bypass surgery is performed.<sup>32</sup> During the demonstration, the bundled rates were updated annually using the payment updates for hospitals and physicians in Medicare's fee-for-service program. According to the government's estimates before the demonstration, the expected savings at the four original hospitals for services covered under the bundled-payment arrangement varied from about 10 percent to 30 percent, with an average of about 20 percent.<sup>33</sup> For each of the three later-starting hospitals, the expected savings were about 7 percent to 8 percent.

Competitive pressures prompted the hospitals to apply for the demonstration. All of the hospitals that applied anticipated that being named a Medicare Participating Heart Bypass Center could help boost their volume of bypass surgeries. The hospitals also were concerned that Medicare could implement a national policy requiring bundled payments for bypass surgery and contract with a limited number of providers; they expected that their participation in the demonstration might improve their chances of inclusion in such a national program.

**Implementation.** The hospitals received the bundled payments and were responsible for paying physicians according to methods they had agreed upon in developing their applications for the demonstration. Each hospital paid a fixed amount per bypass patient to each of the four specialists involved in every bypass case (the thoracic surgeon, cardiologist, anesthesiologist, and radiologist). The hospitals withheld a portion of those fixed payments—typically about 5

<sup>32</sup> The demonstration's two DRGs were 106 (bypass surgery with catheterization) and 107 (bypass surgery without catheterization). It did not include other DRGs in which bypass surgery is performed in addition to other procedures.

<sup>33</sup> Those estimates reflect the estimated savings for the first year of the demonstration and are a weighted average of the savings estimated separately for the two DRGs for which hospitals received separate bundled payments. For each hospital, I weighted the estimated savings for each DRG by the number of admissions in that DRG in the first year of the demonstration.

percent—to establish a pool from which to pay other physicians who sometimes treat bypass patients (such as pulmonologists, nephrologists, and neurologists). The four capitated specialties were at risk for that pool. Specifically, if payments from the pool exceeded the amount that had been set aside, the capitated payments for the four main specialties were reduced. Conversely, if payments from the pool were less than the amount that had been set aside, the four main specialties shared the surplus. The arrangement created an incentive for the four main specialists to economize on referrals to other specialists. One hospital also paid cash bonuses to surgeons who met certain quality-of-care standards and had been judged as reducing the hospital's costs.

Medicare bypass patients faced lower charges for cost sharing under the demonstration than they would have otherwise. After meeting their deductibles, patients treated under the demonstration made a fixed copayment that was designed to be lower than the average coinsurance payments typically paid by bypass patients in their market areas. However, because the great majority of Medicare beneficiaries have supplemental insurance that covers their coinsurance, they did not have a financial incentive to have their surgeries performed at the demonstration hospitals. The demonstration hospitals proposed that they be allowed to waive cost sharing for beneficiaries without supplemental insurance. CMS denied that request, ruling that the hospitals could waive cost sharing for all Medicare patients treated under the demonstration, but not just for those without supplemental insurance. Given that requirement, none of the participating hospitals chose to waive cost sharing under the demonstration.

Under the demonstration, hospitals submitted all of the physician and hospital claims for each heart bypass surgery to the CMS central office in Baltimore, Maryland, rather than to the contractors that process Medicare claims. The hospitals generally received timely payment from CMS for Medicare's portion of the bundled payment, but they experienced major delays in collecting the cost-sharing amounts from supplemental insurers because the cost-sharing arrangements under the demonstration and the documentation that the supplemental insurers received (indicating that Medicare had made a bundled payment and that the beneficiary owed a fixed copayment for all hospital and physician services provided for the surgery) were not compatible with their computer systems.

**Key Findings.** The independent evaluation estimated the effects of the demonstration on Medicare expenditures, examined trends in costs at the participating hospitals, and analyzed patient outcomes under the demonstration.

*Effects on Medicare Expenditures.* The evaluation showed that the demonstration reduced Medicare expenditures on services furnished during hospital stays for bypass surgery by about 10 percent (see Table 3).<sup>34</sup> The estimated savings were in the range of about 5 percent to

---

<sup>34</sup> I derived this estimate from estimates of actual and projected expenditures for services provided during hospital stays for each hospital and year of the demonstration. See Jerry Cromwell and others, *Medicare Participating Heart Bypass Center Demonstration: Volume I, Final Report* (report prepared by Health Economics Research for the Health Care Financing Administration, July 1998), Tables 5-2 and 5-3.

**Table 3.**

**Effects of the Medicare Participating Heart Bypass Center Demonstration on Medicare Expenditures**

Hospital	Average Annual Number of Medicare Heart Bypass Cases During the Demonstration	Percentage Effect on Medicare Expenditures for Services Covered by the Bundled Payment
<b>Four Original Hospitals</b>		
Ohio State University Hospital (Columbus)	157	-22
St. Joseph's Hospital (Atlanta)	745	-8
St. Joseph Mercy Hospital (Ann Arbor)	397	-9
University Hospital (Boston)	251	-19
<b>Three Expansion Hospitals</b>		
Methodist Hospital (Indianapolis)	322	-6
St. Luke's Hospital (Houston)	573	-7
St. Vincent Hospital (Portland, Ore.)	446	-6
<b>All Hospitals</b>	<b>413</b>	<b>-10</b>

Source: Author's analysis of information in Jerry Cromwell and others, *Medicare Participating Heart Bypass Center Demonstration, Volume I, Final Report* (report prepared by Health Economics Research for the Health Care Financing Administration, July 1998).

10 percent for five of the seven hospitals and about 20 percent for the other two hospitals.<sup>35</sup> Those savings reflect the estimated difference between the bundled payments and the amounts that Medicare would have spent for services provided to those bypass patients in the absence of the demonstration.

The evaluation used different approaches to estimate the Medicare payments that would have been made to hospitals in the absence of the demonstration and the payments that would have been made to physicians. For services provided by hospitals, the evaluation used the DRGs on the hospital claims for patients treated under the demonstration to estimate the Medicare payments that would have been made for those beneficiaries under the regular fee-for-service payment rules.<sup>36</sup> Outlier payments that would have been made in the absence of the demonstration were not determined from the claims because the bundled payments were expected to induce hospitals to reduce their costs. Instead, the estimated payments that would have been made were adjusted upward to incorporate the average outlier payments for bypass patients. The evaluation estimated the payments that would have been made for physicians' services furnished to bypass patients during their hospital stays in the absence of the

<sup>35</sup> The evaluation did not provide information on the statistical precision of those estimates. Under reasonable assumptions about the variance of Medicare expenditures for bypass surgery, however, it appears that all of the estimates are significantly different from zero at the 5 percent level.

<sup>36</sup> Under the demonstration, hospitals were required to submit special Medicare claims for their services and those of the physicians treating bypass patients. The claims contained all information required for claims submitted in the fee-for-service Medicare program. They were not used to establish payment amounts but rather to identify the bypass surgeries for which hospitals and physicians were to be paid the bundled rates and to provide data for the evaluation.

demonstration by defining a standard package of physicians' services associated with bypass surgery (based on an analysis of national Medicare claims) and adjusting the cost of that package to account for differences in payment rates in different areas of the country and at different points in time.<sup>37</sup> An important limitation of that approach is that the mix of physicians' services that would have been used by heart bypass patients at the demonstration hospitals in the absence of the demonstration may have differed from the standard package of services assumed for the analysis because of differences among hospitals in terms of physicians' practice styles and patients' severity of illness and changes in physicians' practice patterns over time.

The evaluation also estimated the effects of the demonstration on Medicare expenditures by a broader measure that included the effects just described along with two additional components: effects on expenditures over the 90 days after discharge from the hospital and effects on expenditures resulting from any change in the demonstration hospitals' share of Medicare bypass surgeries in their local market areas. That analysis estimated that the two additional components may have accounted for relatively small savings, but given the challenges in estimating the effects, the estimates are highly uncertain. The demonstration also could have caused some physicians to attempt to steer less complex cases to the demonstration hospitals and more complex cases to other hospitals, but the evaluation did not examine that question.

The demonstration could have affected Medicare's expenditures after discharge from the hospital in various ways. On the one hand, the incentives created by the bundled-payment system to economize on the use of services during a hospital stay could have resulted in some patients' being discharged earlier and in poorer health, with greater need for follow-up care. On the other hand, to the extent that providers responded to those incentives by increasing cooperation and coordination, quality of care could have improved under the demonstration, resulting in patients' requiring less care after discharge. Moreover, the incentive to reduce readmissions under bundled payments also could have led to improved quality of care and less need for care following discharge.

To estimate the effects of the demonstration on Medicare expenditures during the 90 days after patients are discharged, the evaluation compared actual postdischarge expenditures for beneficiaries treated under the demonstration with the estimated expenditures that would have been incurred for those beneficiaries in the absence of the demonstration. That estimate was derived by calculating Medicare's average postdischarge expenditures for heart bypass patients treated at the demonstration hospitals in the year before the demonstration and increasing those estimates to reflect national growth in postdischarge Medicare payments per bypass patient from the base year to each year of the demonstration. The estimated effects of the demonstration on postdischarge expenditures varied among hospitals and were imprecise; the point estimate indicated that the demonstration reduced postdischarge expenditures by about 4 percent. (The

---

<sup>37</sup> The evaluation did not use the special claims for physicians' services submitted for the demonstration to estimate the expenditures that would have been incurred for those services in the absence of the demonstration because the bundled-payment system gave physicians incentives to economize on the use of resources (such as visits by consulting specialists). Consequently, the claims probably would have underestimated the expenditures for physicians' services that would have been incurred in the absence of the demonstration.

precision of the effects of the demonstration on other measures of Medicare expenditures was not assessed.)

The approach used to estimate effects on postdischarge expenditures has several important limitations. Using a single year's worth of data on postdischarge expenditures for each hospital before the demonstration as the base year for estimating the postdischarge expenditures that would have been observed in the absence of the demonstration introduced significant uncertainty because average postdischarge expenditures can vary greatly from year to year for samples of the size that were used in the analysis. (On average, about 400 Medicare bypass patients were treated at each participating hospital in the year before the demonstration.) In addition, the national trends in per capita postdischarge expenditures for bypass patients that were used to update the base year expenditures may have differed from the trends that would have been observed for demonstration patients in the absence of the demonstration because of differences in factors such as physicians' practice styles, the quality of care delivered, and patients' severity of illness. To test the sensitivity of the results to the use of national growth rates to update the base year expenditures, the evaluation also used update factors based on the change in postdischarge expenditures for bypass patients at other hospitals in the market areas of the demonstration hospitals. Using those local update factors yielded somewhat higher estimated savings. Those estimates suffer from the same limitations as those generated from the national update factors, however, because of possible differences between the demonstration hospitals and other local hospitals in physicians' practice styles, quality of care, and patients' severity of illness. The evaluators responded to the considerable uncertainties surrounding those estimates by reporting the lower savings estimates generated using the national update factors as the preferred estimates.

The third component of the estimated effect of the demonstration on Medicare's expenditures captured any changes in expenditures that were attributable to changes in the demonstration hospitals' share of Medicare bypass surgeries in their local market areas. If the participating hospitals increased their market shares as a result of the demonstration, the demonstration would have reduced Medicare spending if the average Medicare payment per bypass surgery was lower in the demonstration hospitals than in the hospitals where those patients would otherwise have been treated. The evaluation defined the market area for each demonstration hospital to include all hospitals located in its metropolitan area plus any additional hospitals that the demonstration hospital regarded as a competitor for bypass surgeries. The demonstration hospitals' market shares did not exhibit any clear trends during the demonstration; some hospitals had slightly higher market shares at the end of the demonstration than they had in the year before it started, and others had slightly lower market shares. Most hospitals' market shares fluctuated slightly during the demonstration. On balance, the evaluation showed that those changes in market share yielded relatively small savings for Medicare.

That estimate of Medicare savings attributable to shifts in market share is highly uncertain and should be viewed with caution. Any changes in the hospitals' market shares could have resulted from any of a variety of factors other than the demonstration, such as random fluctuations in the incidence of severe heart conditions in the communities that serve as primary referral areas for different hospitals and actions by hospitals to increase or reduce their volume of heart bypass surgeries. Indeed, the hospitals that participated in the demonstration all wanted to increase their volume of bypass surgeries, and they might have done so without the demonstration. In addition,

the evaluation used each demonstration hospital's market share in the year before the demonstration to estimate what its share would have been in subsequent years in the absence of the demonstration. Using a single year as a base period for such an analysis is an important limitation because market share can change over time for a variety of reasons.

*Trends in the Demonstration Hospitals' Costs.* Three of the four hospitals originally in the demonstration reduced their costs for treating Medicare bypass patients during the demonstration. From 1990 (the year before the demonstration) to 1993, those three hospitals reduced their average cost per bypass patient by 7 percent to 23 percent for cases in one DRG and by 2 percent to 13 percent for cases in another DRG. The reduction in costs was attributable primarily to lower costs for general nursing and intensive care, pharmacy, and laboratory services. At the fourth hospital, costs rose by 11 percent for one DRG and by 24 percent for the other DRG. Those estimated changes in costs were not adjusted for inflation and only include costs incurred by the hospitals, not the physicians. During that period, the costs of hospitals' inputs rose nationally by about 11 percent.<sup>38</sup>

The analysis was based on detailed cost data submitted by the four original demonstration hospitals for 1990 through 1993. Comparable data were not available for other hospitals, so it is not known how the changes in costs for Medicare bypass surgeries at the four demonstration hospitals compared with changes in costs at other hospitals during that period. However, an important factor that influences hospitals' costs—patients' length of stay—declined nationally for Medicare bypass patients by 18 percent (from 15.0 days to 12.3 days) between 1990 and 1993. That suggests that significant changes in the management of such patients were occurring nationally during that period. By comparison, the average length of stay for Medicare bypass patients at the four original demonstration hospitals fell during that period by 14 percent to 32 percent, with a median reduction of 26 percent.<sup>39</sup> It is not possible to determine from those estimates whether the demonstration reduced lengths of stay because the differences between the trends under the demonstration and national trends may be due in part to differences between hospitals in various other factors that influence lengths of stay.

Interviews with hospital managers and clinicians at the demonstration hospitals identified some important changes in their approaches to patient management that may have contributed to the lower costs. They included greater involvement by surgeons in postoperative care, earlier discharge of patients from the intensive care unit, greater standardization of surgical protocols and supplies, substitution of less expensive drugs for more costly ones, and greater reliance on critical nurse specialists for managing patients' care in the hospital. Because comparable information was not collected from other hospitals, it is not known how those changes in patient management at the demonstration hospitals compare with changes that may have been occurring nationwide.

---

<sup>38</sup> That estimated increase in hospital input costs is based on growth in the market basket index that CMS uses to update Medicare payment rates for inpatient hospital care.

<sup>39</sup> The estimates of lengths of stay for Medicare heart bypass patients for the demonstration and nationally are based on information in Jerry Cromwell and others, *Medicare Participating Heart Bypass Center Demonstration: Volume I, Final Report*. The estimates account for only the hospital stays in the two DRGs that were included in the demonstration. To standardize comparisons, the evaluation weighted each hospital's average length of stay for each DRG by the national proportion of cases in each DRG.

*Patient Outcomes.* During the demonstration, the health outcomes of bypass patients treated at the participating hospitals were similar to those of bypass patients treated at the comparison hospitals in the local market areas. In particular, the two groups of patients had similar in-hospital mortality rates, complication rates, and postsurgical health and functional status. In addition, the in-hospital mortality rate for Medicare heart bypass patients declined at the same rate during the demonstration at the demonstration hospitals and the comparison hospitals. For some of those outcomes, the evaluation controlled for differences between demonstration patients and the comparison group on demographic characteristics and measures of health status before the surgery. Those results suggest that the bundled payment did not result in deterioration in patient outcomes. However, the evaluation results do not provide a definitive estimate of the effect of bundled payment on health outcomes because there may have been preexisting differences between the demonstration hospitals and the comparison hospitals that may have resulted in differences in patient outcomes in the absence of the demonstration.

## Strengths and Weaknesses of the Demonstration Designs

In assessing the strengths and weaknesses of the demonstration designs, it is useful to distinguish design features that influenced the ability of the evaluations to measure the effects of the demonstrations from issues regarding the manner in which the interventions were structured and implemented.

### Measurement Issues

A key strength of the Home Health Pay-for-Performance Demonstration was the randomized design. None of the other three demonstrations of value-based payment used a randomized design. Such a design would not have been feasible for the PGP demonstration or the Heart Bypass demonstration, given the relatively small number of organizations that participated and the fact that the interventions were organization-wide. The design of the Premier Hospital Quality Incentive Demonstration would have been stronger if the 278 hospitals that applied to participate had been randomly assigned either to a treatment group or to a control group. However, the demonstration arose from an unsolicited proposal by Premier and the hospitals probably would have been less likely to participate if they had known that some would be randomly assigned to a control group.

The three demonstrations that did not use a randomized design took different approaches to specifying a comparison standard for estimating the effects of the demonstrations. The PGP demonstration estimated the Medicare expenditures that would have been incurred for the PGPs' patients in the absence of the demonstration by first computing the average expenditure for the patients of each PGP during a base year before the demonstration and then increasing that amount by the growth in the average expenditure for a comparison group of beneficiaries from the PGP's local service area. This approach thus used the growth in average expenditures for the comparison group to estimate the expenditure growth that would have been observed for the PGP patients in the absence of the demonstration. Although it was a reasonable approach under the circumstances, there was a considerable risk that the comparison groups differed from the PGP patients in ways that affected their cost growth. Physicians at the large group practices that volunteered and were selected for the PGP demonstration may differ from other physicians in the community in their practice styles, and beneficiaries who obtain care from the two groups of

providers may differ in terms of health status and attitudes toward care—and those differences could result in differences in cost growth. Indeed, the evaluation showed that cost growth before the demonstration was lower for the patients of the four PGP s that received bonuses than it was for their comparison groups, suggesting that the savings estimated for those PGP s might be overstated.

Because of concerns about using local comparison groups, CMS is using figures on the national growth in Medicare spending—rather than growth in spending by local comparison groups—to establish spending targets for the PGP Transition Demonstration, which began in January 2011. That approach also has limitations, however; the national growth in spending may not be a good measure of the spending growth that would have occurred for the PGP s' patients absent the demonstration. The challenge in setting an expenditure target is not unique to the PGP demonstration but would apply to any effort to implement the payment approach more broadly.

The evaluation of the Premier demonstration estimated demonstration impacts by comparing expenditures and quality of care for Medicare patients at the demonstration hospitals with corresponding outcomes for Medicare patients at other hospitals. The statistical analysis attempted to remove the effects of differences in the characteristics of the hospitals, their market areas, and their patients. As noted previously, however, such designs carry the risk that differences in outcomes that result from underlying differences between the demonstration hospitals and the comparison hospitals will be confounded with the demonstration impacts.

The Heart Bypass demonstration used different comparison standards for estimating the effects of the demonstration on different outcome measures. Those different comparison standards, and their strengths and limitations, were discussed previously.

### **Design Issues**

An inherent limitation of the payment approach tested in the PGP demonstration is that the PGP s were accountable for managing the care of beneficiaries who retained complete freedom to obtain care from other providers in a fee-for-service environment. Moreover, the PGP s did not know which particular beneficiaries they were accountable for in a given year until sometime in the next year.

Another challenge facing the PGP demonstration was providing the participating physician groups with performance feedback and bonuses promptly. CMS provided the participating physician groups with feedback on their performance during a given year approximately 12 months after the end of the year and distributed bonus payments about 3 months after that.<sup>40</sup> Determining bonus payments was time consuming because beneficiaries were retroactively assigned to physician groups at the end of each year based on their claims experience. CMS allowed six months, to ensure that all claims had been submitted for the year, and then required additional time to assign beneficiaries to groups and analyze data on their costs and quality of care. Some physician groups noted that being provided feedback sooner would have enabled

---

<sup>40</sup> Government Accountability Office, *Medicare Physician Payment: Care Coordination Programs Used in Demonstration Show Promise, but Wider Use of Payment Approach May Be Limited*, GAO-08-65 (February 2008), [www.gao.gov/new.items/d0865.pdf](http://www.gao.gov/new.items/d0865.pdf).

them to make adjustments to their programs. Although CMS provided the physician groups with quarterly claims data on beneficiaries they had treated, some groups did not have the resources to analyze those data.

A key design decision for the PGP demonstration was the share of any estimated Medicare savings that the PGPs would receive. Setting that parameter requires balancing two competing objectives. On the one hand, the greater the share of estimated savings that goes to the PGPs, the stronger is their incentive to implement changes in care processes that would reduce costs. On the other hand, for a given amount of savings, the greater the share that goes to the PGPs, the less is available to be captured by the government. The shared-savings arrangement specified for the demonstration, in which the PGPs received up to 80 percent of any savings, gave the PGPs an incentive to reduce costs and also allowed the government to share in any savings.

A key design decision for the Premier Hospital Quality Incentive Demonstration was the structure of the financial incentives offered to hospitals. The first three years of the demonstration (which was the period covered by the evaluation) tested only one approach to specifying those financial incentives: Bonuses went to hospitals whose quality scores were in the top 20 percent, but hospitals were not rewarded based on improvement. The lack of any reward for improvement, combined with the fact that hospitals were permitted to withdraw from the demonstration before penalties for low scores went into effect (in year 3), limited the incentive for many hospitals to make changes in their care processes to try to improve their scores.

A limitation of the Home Health Pay-for-Performance Demonstration is that, unlike the PGP demonstration, there is no threshold percentage by which actual spending is required to fall below projected spending in order to be considered savings. In the first year of the demonstration, CMS returned all of the estimated savings to some of the HHAs in three regions, but the small amount of the savings (ranging from one region to another by 0.4 percent to 1.2 percent) suggests that the differences between actual and projected expenditures in those regions might reflect random variation in Medicare expenditures rather than actual savings. Moreover, the “negative savings” in one region were not used to offset the estimated savings in the other three regions. Because of those features, it is likely that the demonstration resulted in a small net increase in Medicare expenditures. The basic design of this demonstration was successful in encouraging large numbers of HHAs to participate, but attempts to expand such a payment approach in the future will not reduce Medicare expenditures unless CMS establishes a threshold for determining whether HHAs achieve savings and then shares those savings with the HHAs.

## **Discussion**

The PGP demonstration had little or no net effect on Medicare expenditures, after accounting for the bonuses that were paid to the physician groups. Moreover, the amount of savings estimated for that demonstration might be overstated because the Centers for Medicare & Medicaid Services (CMS) overestimated the expenditures that the PGPs’ patients would have incurred in the absence of the demonstration. The Premier Hospital Quality Incentive Demonstration also had no discernible effect on Medicare expenditures. The evaluations found that those two demonstrations improved the quality of care (as reflected by the particular measures adopted for the demonstrations) by a small amount. Preliminary results for the Home Health Pay-for-Performance Demonstration indicate that it had little or no effect on Medicare expenditures or

quality of care in the first year. The apparent lack of success of the PGP demonstration at reducing Medicare expenditures is particularly noteworthy because all of the participating organizations were large multispecialty group practices that had experience with care management for other populations before the demonstration. The results from that demonstration suggest that additional experimentation, evaluation, and refinement over a period of years will probably be needed to identify approaches involving pay-for-performance and shared-savings models that can significantly reduce Medicare expenditures.

The Heart Bypass demonstration alone among the demonstrations discussed here yielded significant savings to the Medicare program. The evaluation showed that bundled payment reduced Medicare expenditures on heart bypass surgeries by about 10 percent and that there were no apparent adverse effects on patient outcomes. The savings were generated because Medicare was able to negotiate bundled payment rates with the seven hospitals and the physicians on their medical staffs that were lower than separate payments would have been.

Medicare operated a bundled-payment demonstration for cataract surgery at approximately the same time.<sup>41</sup> That demonstration is not discussed in detail here because a full evaluation was not completed. The demonstration was actively opposed by national and local societies of ophthalmologists, who filed suit in federal court seeking to prevent its implementation. The government prevailed, however, and implemented the demonstration with four provider groups in three metropolitan areas. Given the bundled-payment rates that were negotiated with those providers, the government projected that it would realize savings on cataract surgeries of 2 percent to 5 percent. A formal evaluation estimating the effects of the demonstration on Medicare expenditures was not conducted, however.

After the savings in the Heart Bypass demonstration were reported, Medicare began in 1995 to develop another demonstration to test bundled payment for certain cardiovascular and orthopedic procedures.<sup>42</sup> For various reasons, that demonstration was never implemented.<sup>43</sup> However, another demonstration of bundled payment for cardiovascular and orthopedic procedures—the Acute Care Episode (ACE) Demonstration—began in 2009 at five health systems. Evaluation results are not yet available, but based on the bundled payment rates, CMS has projected that Medicare savings will range from 1 percent to 6 percent, depending on the provider and type of procedure.<sup>44</sup> Appendix B discusses the ACE demonstration and other ongoing demonstrations of value-based payment for which evaluation results are not available.

---

<sup>41</sup> See Abt Associates, *Medicare Cataract Surgery Alternate Payment Demonstration: Final Evaluation Report* (prepared for the Health Care Financing Administration, June 1997).

<sup>42</sup> See Pierre L. Yong, Robert S. Saunders, and LeighAnne Olsen, eds., *The Healthcare Imperative: Lowering Costs and Improving Outcomes. Workshop Series Summary* (Washington, D.C.: National Academies Press, 2010) Chapter 15, pp. 473–492, [www.nap.edu/openbook.php?record\\_id=12750](http://www.nap.edu/openbook.php?record_id=12750).

<sup>43</sup> The demonstration was not implemented because of competing resource requirements for implementing the provisions of the Balanced Budget Act of 1997 and a moratorium on computer system changes because of the “year 2000 problem.” In addition, the demonstration faced organized opposition from the American Academy of Orthopedic Surgeons.

<sup>44</sup> Yong, Saunders, and Olsen, *The Healthcare Imperative: Lowering Costs and Improving Outcomes*, pp. 473–492.

The results of the Heart Bypass demonstration suggest that bundled payment holds promise as an approach for reducing Medicare expenditures. That demonstration yielded savings because the bundled payments that the seven hospitals and the physicians on their medical staffs agreed to accept for treating Medicare heart bypass patients were lower than the sum of the separate payments that they otherwise would have received from Medicare. The amount of savings that can be generated from similar initiatives in the future will depend on the willingness of hospitals and physicians to accept such terms.

## Appendix A

# Studies of the Medicare Demonstrations Reviewed for This Paper

### Physician Group Practice Demonstration

Department of Health and Human Services, *Report to Congress: Physician Group Practice Demonstration Evaluation Report* (2009), [www.cms.gov/reports/downloads/RTC\\_Sebelius\\_09\\_2009.pdf](http://www.cms.gov/reports/downloads/RTC_Sebelius_09_2009.pdf).

Department of Health and Human Services, *Report to Congress: Physician Group Practice Demonstration First Evaluation Report* (2006), [www.cms.gov/reports/downloads/Leavitt1.pdf](http://www.cms.gov/reports/downloads/Leavitt1.pdf).

John Kautter and others, "Medicare Physician Group Practice Demonstration Design: Quality and Efficiency Pay-for-Performance," *Health Care Financing Review*, vol. 29, no. 1 (Fall 2007), pp. 15–29.

Michael Trisolini and others, *The Medicare Physician Group Practice Demonstration: Lessons Learned on Improving Quality and Efficiency in Health Care* (paper prepared by RTI International for the Commonwealth Fund, February 2008), <http://bit.ly/z79VJL>.

Michael Trisolini and others, *Medicare Physician Group Practices: Innovations in Quality and Efficiency* (paper prepared by RTI International for the Commonwealth Fund, December 2006), [www.commonwealthfund.org/usr\\_doc/971\\_Trisolini\\_Medicare\\_physician\\_group\\_practices\\_i.pdf](http://www.commonwealthfund.org/usr_doc/971_Trisolini_Medicare_physician_group_practices_i.pdf).

### Premier Hospital Quality Incentive Demonstration

Centers for Medicare & Medicaid Services, *Evaluation of the Premier Hospital Quality Incentive Demonstration. Executive Summary: Impacts on Quality of Care, Medicare Reimbursements, and Medicare Beneficiaries' Length of Stay During the First Three Years of the Demonstration* (March 2009), [www.cms.gov/reports/downloads/Premier\\_ExecSum\\_2010.pdf](http://www.cms.gov/reports/downloads/Premier_ExecSum_2010.pdf).

Stephen Kennedy and others, *Evaluation of the Premier Hospital Quality Incentive Demonstration. Impacts on Quality, Medicare Inpatient Reimbursements, and Medicare Lengths of Stay: Volumes I and II* (report submitted by Abt Associates to the Centers for Medicare & Medicaid Services, December 2008), Executive summary online: [www.cms.gov/reports/downloads/Premier\\_ExecSum\\_2010.pdf](http://www.cms.gov/reports/downloads/Premier_ExecSum_2010.pdf).

Andrew M. Ryan, "Effects of the Premier Hospital Quality Incentive Demonstration on Medicare Patient Mortality and Cost," *Health Services Research*, vol. 44, no. 3 (June 2009), pp. 821–842.

Peter K. Lindenauer and others, "Public Reporting and Pay for Performance in Hospital Quality Improvement," *New England Journal of Medicine*, vol. 356, no. 5 (February 1, 2007), pp. 486–496.

### Medicare Participating Heart Bypass Center Demonstration

Jerry Cromwell and others, *Medicare Participating Heart Bypass Center Demonstration: Volume I, Final Report* (report prepared by Health Economics Research for the Health Care Financing Administration, July 1998), [www.cms.gov/demoprojectsevalrpts/MD/ItemDetail.asp?ItemID=CMS063472](http://www.cms.gov/demoprojectsevalrpts/MD/ItemDetail.asp?ItemID=CMS063472).

Jerry Cromwell and others, *Medicare Participating Heart Bypass Center Demonstration: Volume II, Final Report* (report prepared by Health Economics Research for the Health Care Financing Administration, July 1998), [www.cms.gov/demoprojectsevalrpts/MD/ItemDetail.asp?ItemID=CMS063472](http://www.cms.gov/demoprojectsevalrpts/MD/ItemDetail.asp?ItemID=CMS063472).

Jerry Cromwell, Debra A. Dayhoff, and Armen H. Thoumaian. "Cost Savings and Physician Responses to Global Bundled Payments for Medicare Heart Bypass Surgery," *Health Care Financing Review*, vol. 19, no. 1 (Fall 1997), pp. 41–57.

### **Medicare Home Health Pay-for-Performance Demonstration**

David F. Hittle, Eugene J. Nuccio, and Angela A. Richard, *Evaluation of the Medicare Home Health Pay-for-Performance Demonstration. CY2008 Report—Volume 1: Agency Characteristics, Costs, and Quality Measure Performance Among Treatment, Control, and Non-Participant Groups* (report prepared by Division of Health Care Policy and Research, University of Colorado Anschutz Medical Campus for the Centers for Medicare & Medicaid Services, June 2011),  
[www.cms.gov/reports/downloads/HHP4P\\_Demo\\_Eval\\_2008\\_Vol1.pdf](http://www.cms.gov/reports/downloads/HHP4P_Demo_Eval_2008_Vol1.pdf).

## **Appendix B**

### **Other Demonstrations for Which Results Are Not Yet Available**

Four other Medicare demonstrations of value-based payment are under way, but evaluations of their effects on Medicare expenditures or other outcomes are not available. The Medicare Care Management Performance Demonstration is a three-year project involving approximately 640 small and medium-sized primary care physician practices in four states. Initially, the physician practices received a payment from the Centers for Medicare & Medicaid Services (CMS) for reporting baseline data on 26 quality-of-care measures for the year before the demonstration. In each year of the demonstration, the practices are eligible to receive a bonus if their quality scores exceed an established threshold. The measures concern the care of patients with congestive heart failure, coronary artery disease, and diabetes, and the provision of preventive care. The practices receive an additional payment from CMS if they report their data electronically from an electronic health record (EHR) system that meets certain industry standards.

The demonstration was mandated by the Medicare Prescription Drug, Improvement, and Modernization Act of 2003 (MMA). It began in 2007 and is scheduled to last three years. A randomized design was not used for the demonstration. Effects on Medicare expenditures, quality of care, and health will be estimated by comparing outcomes for patients of the demonstration physician practices with outcomes for patients of similar practices in other geographic areas.

The Electronic Health Records Demonstration was initiated by CMS and is similar in some respects to the Medicare Care Management Performance Demonstration. Small and medium-sized primary care physician practices in four geographic areas that applied to participate and met the eligibility criteria were randomly assigned to a treatment group or a control group. In each year of the demonstration, physician practices in the treatment group are eligible to receive a bonus for using an EHR with certain core minimum functionalities. In the second year, the physician practices in the treatment group will receive a bonus for reporting data on quality of care measures. Beginning in the third year, they will be eligible to receive a bonus if their quality scores exceed a certain threshold. The demonstration began in 2009 and will last five years. Effects on Medicare expenditures, quality of care, and health will be estimated by comparing outcomes for patients of the demonstration physician practices with outcomes for patients of practices in the control group.

The Medicare Health Care Quality Demonstration Program, which was mandated by the MMA, gives the Secretary of Health and Human Services broad authority to develop demonstration projects aimed at improving the quality of care and lowering costs. Two large regional organizations began demonstrations under that authority in 2010: the Indiana Health Information Exchange and North Carolina Community Care Networks. In each demonstration, physicians will be eligible to receive shared-savings bonuses if they meet certain targets regarding quality of care and Medicare expenditures.

The Acute Care Episode Demonstration was initiated by CMS to test bundled payments that cover all physician and hospital services for selected cardiac and orthopedic procedures. The demonstration began in 2009 at five health systems; one of them is accepting bundled payment

for orthopedic procedures, two are accepting bundled payment for cardiac procedures, and two are accepting bundled payment for both. Evaluation results are not yet available for the demonstration, but based on the bundled-payment rates, CMS has projected that Medicare savings will range from 1 percent to 6 percent, depending on the provider and type of procedure.<sup>45</sup>

---

<sup>45</sup> Pierre L. Yong, Robert S. Saunders, and LeighAnne Olsen, eds., *The Healthcare Imperative: Lowering Costs and Improving Outcomes. Workshop Series Summary* (Washington, D.C.: National Academies Press, 2010), Chapter 15, pp. 473–492, [www.nap.edu/openbook.php?record\\_id=12750](http://www.nap.edu/openbook.php?record_id=12750).