2018

Evaluation of medicare's bundled payments initiative for medical conditions

Karen E. Joynt Maddox
Washington University School of Medicine in St. Louis

E. John Orav
Brigham and Women's Hospital

Jie Zheng
Harvard University

Arnold M. Epstein
Harvard University

SPECIAL ARTICLE

Evaluation of Medicare's Bundled Payments Initiative for Medical Conditions

Karen E. Joynt Maddox, M.D., M.P.H., E. John Orav, Ph.D., Jie Zheng, Ph.D., and Arnold M. Epstein, M.D.

ABSTRACT

BACKGROUND

The Center for Medicare and Medicaid Innovation (CMMI) launched the Bundled Payments for Care Improvement (BPCI) initiative in 2013. A subsequent study showed that the initiative was associated with reductions in Medicare payments for total joint replacement, but little is known about the effect of BPCI on medical conditions.

METHODS

We used Medicare claims from 2013 through 2015 to identify admissions for the five most commonly selected medical conditions in BPCI: congestive heart failure (CHF), pneumonia, chronic obstructive pulmonary disease (COPD), sepsis, and acute myocardial infarction (AMI). We used difference-in-differences analyses to assess changes in standardized Medicare payments per episode of care (defined as the hospitalization plus 90 days after discharge) for these conditions at BPCI hospitals and matched control hospitals.

RESULTS

A total of 125 hospitals participated in BPCI for CHF, 105 hospitals for pneumonia, 101 hospitals for COPD, 88 hospitals for sepsis, and 73 hospitals for AMI. At baseline, the average Medicare payment per episode of care across the five conditions at BPCI hospitals was $24,280, which decreased to $23,993 during the intervention period (difference, −$286; P=0.41). Control hospitals had an average payment for all episodes of $23,901, which decreased to $23,503 during the intervention period (difference, −$398; P=0.08; difference in differences, $112; P=0.79). Changes from baseline to the intervention period in clinical complexity, length of stay, emergency department use or readmission within 30 or 90 days after hospital discharge, or death within 30 or 90 days after admission did not differ significantly between the intervention and control hospitals.

CONCLUSIONS

Hospital participation in five common medical bundles under BPCI was not associated with significant changes in Medicare payments, clinical complexity, length of stay, emergency department use, hospital readmission, or mortality. (Funded by the Commonwealth Fund.)

From Washington University School of Medicine, St. Louis (K.E.J.M.); and Brigham and Women's Hospital (E.J.O., A.M.E.) and Harvard T.H. Chan School of Public Health (E.J.O., J.Z., A.M.E.) — both in Boston. Address reprint requests to Dr. Joynt Maddox at Washington University School of Medicine, 660 S. Euclid Ave, St. Louis, MO 63130, or at kjoyntmaddox@wustl.edu.

N Engl J Med 2018;379:260-9.
DOI: 10.1056/NEJMsa1801569

Episode-based payment holds promise for improving the quality and efficiency of care. The Center for Medicare and Medicaid Innovation (CMMI) launched the Bundled Payments for Care Improvement (BPCI) initiative in 2013.[1] BPCI is a voluntary program, and hospitals that choose to participate may select from several models. Under model 2, the most popular model among hospitals, participating hospitals assume accountability for the costs of all care within 30, 60, or 90 days after hospitalization for 1 or more of 48 conditions that together account for approximately 70% of Medicare spending.[2] Hospitals may choose the conditions and the time window. If cost targets are achieved, participating hospitals keep a portion of the savings; if cost targets are exceeded, participating hospitals reimburse Medicare for part of the difference.

BPCI is a large demonstration project, with hundreds of participating hospitals, and has enjoyed bipartisan support. The Department of Health and Human Services has announced that a closely related version of the project (BPCI Advanced) will be launched at the beginning of fiscal year 2019.[3] Despite the importance of episode-based payment, however, there has been little research examining its efficacy or determining whether it has unintended consequences, such as hospitals' selecting patients with relatively less complex conditions to reduce costs and improve outcomes. A previous peer-reviewed evaluation of the program,[4] which was restricted to patients hospitalized for joint replacement, showed a modest reduction in Medicare-allowed payments, no decrease in the complexity of cases treated at BPCI hospitals, and no change in mortality. Evaluations performed by a federal contractor have had similar findings.[5,6] Because BPCI and BPCI Advanced include a wide variety of clinical conditions, there is a critical need for information on how bundled-payment models might change patterns of care for medical conditions.

We aimed to fill this gap by studying the five most common medical bundles under BPCI: congestive heart failure (CHF), pneumonia, chronic obstructive pulmonary disease (COPD), sepsis, and acute myocardial infarction (AMI). Patients with these conditions account for nearly two thirds of patients enrolled in medical bundles in BPCI.

Methods

Overview of BPCI

There are four participation models in BPCI; more than 95% of participating hospitals are in model 2.[1] Under BPCI, hospitals are paid fee-for-service rates, but Medicare-allowed payments are retrospectively reconciled against targets on a quarterly basis. Target prices are based on historical spending, minus a discount of 2 to 3%, depending on the episode length chosen by the hospital.

The Commonwealth Fund, which provided funding for the study, had no role in its design or conduct; the collection, management, analysis, or interpretation of the data; the preparation, review, or approval of the manuscript; or the decision to submit the manuscript for publication. The requirement for informed consent was waived because the data were deidentified. The study was approved by the Office of Human Research Administration at the Harvard T.H. Chan School of Public Health.

BPCI Participants

We obtained from CMMI publicly available lists of participating hospitals in the BPCI initiative, their start date for financial incentives, and the date they planned to terminate participation. From these lists, we identified hospitals enrolled by July 1, 2015, in bundles for CHF, pneumonia, COPD, sepsis, or AMI. These data were linked both to hospital characteristics, obtained from the 2014 American Hospital Association file, and to market characteristics such as availability of post–acute care services and median income levels, obtained from the Area Resource File. For each hospital, market share was calculated as the proportion of all admissions in the county for the condition of interest. Market competitiveness was calculated with the use of the Herfindahl–Hirschman Index[7] (Table 1).

Control Hospitals

To identify control hospitals, we used a modification of the approach used in a previous study of BPCI.[4] With the use of propensity scores based on market and hospital characteristics, as well as baseline rates of discharge to skilled nursing facilities and of readmissions, each hospital participating in BPCI was matched without replacement with up to three control hospitals

Table 1. Baseline Hospital and Market Characteristics.*

Characteristic	BPCI Hospitals (N=492)	Matched Control Hospitals (N=898)	P Value	All Non-BPCI Hospitals (N=3681)	P Value, BPCI vs. All Non-BPCI Hospitals
Participation in bundles					
Mean no. of bundles per hospital	2.6	NA		NA	
≥2 bundles (% of hospitals)	62.6	31.7		NA	
Mean discharges (no. per hospital)	153.3	146.1	0.34	102.2	<0.001
Mean 90-day episode payment in 2013 (U.S. dollars)	23,161	22,664	0.33	21,977	0.02
Type of hospital (%)					
Nonprofit	74.6	76.5	0.43	64.9	<0.001
Public	2.4	3.6	0.25	15.7	<0.001
For-profit	23.0	19.9	0.18	19.4	0.049
Teaching	47.2	48.7	0.59	32.1	<0.001
Urban location (%)	100	100	1.00	91.0	<0.001
Geographic region (%)					
Northeast	28.9	30.2	0.61	15.9	<0.001
Midwest	14.6	16.8	0.29	23.7	<0.001
South	27.2	24.9	0.35	42.4	<0.001
West	29.3	28.1	0.63	18.0	<0.001
Disproportionate-share hospital (%)	29.2	29.7	0.59	29.2	0.92
Mean size (no. of beds)	322	315	0.61	229	<0.001
Market characteristics					
Median household income (U.S. dollars)	56,174	57,002	0.29	51,114	<0.001
Medicare enrollees who are in Medicare Advantage plans (%)	29.7	29.1	0.41	26.5	<0.001
Skilled nursing facilities (total no. of beds)	5922	5963	0.93	3638	<0.001
Hospital market share	0.31	0.32	0.73	0.50	<0.001
HHI	0.14	0.14	0.76	0.13	0.19

* All data on hospital and market characteristics are from 2013. Markets were defined on the basis of counties. HHI denotes Herfindahl–Hirschman index (defined as the sum of the squares of each provider's market share, with a perfectly competitive market having an HHI near zero and a completely concentrated market having an HHI of 1), and NA not applicable.

whose log-odds propensity scores were within 0.515 of the log-odds propensity score for the BPCI hospital. We chose this difference of 0.515 to be consistent with the previous work in this area[4] and because it falls within the recommended difference of 0.84,[8] which is based on the pooled standard deviation of our log-odds propensity scores multiplied by 0.2. Nonparticipating hospitals were excluded if they were not paid through the inpatient prospective-payment system or if they participated in any other bundling model or for any other condition in BPCI.

PATIENT IDENTIFICATION

We used Medicare inpatient files from January 1, 2013, through September 30, 2015, to identify index admissions for CHF (Medicare Severity Diagnosis-Related Groups [MS-DRGs] 291–293),

pneumonia (MS-DRGs 177–179 and 193–195), COPD (MS-DRGs 190–192, 202, and 203), sepsis (MS-DRGs 870–872), and AMI (MS-DRGs 280–282) at participating hospitals and matched control hospitals. We included only beneficiaries who were continuously enrolled in Medicare Parts A and B during their episode of care and excluded those with Medicare eligibility because of end-stage renal disease. Claims contained demographic characteristics, principal discharge diagnoses, coexisting conditions, and service use.

For each index hospitalization, standardized Medicare-allowed payments for episodes of care (defined as the hospitalization plus 90 days after discharge), hereafter referred to as Medicare payments, were calculated with the use of 100% of claims for inpatient care, skilled nursing facility care, home health agency services, and durable medical equipment, as well as a 20% random sample of claims for outpatient services and physician fees. Standardized payments specified by the Centers for Medicare and Medicaid Services (CMS) reflect a process that removes differences in payment according to a wage index, as well as differences with respect to payments for indirect costs of medical education, payments to disproportionate-share hospitals (i.e., hospitals providing a disproportionate share of care to Medicare beneficiaries and uninsured patients), and other special payments. Total payments were Winsorized at the 95th percentile of national episode payments and adjusted for inflation to prices in 2015. Because less than 5% of hospitals chose 30- or 60-day episodes, we analyzed only 90-day episodes, which is consistent with the previous peer-reviewed evaluation.[4]

STUDY PERIOD

We considered baseline to be 9 months to 3 months (a 6-month period) before each hospital's start date, with the intervention period starting immediately after the start date. The intervention period ranged from 3 to 9 months according to the enrollment date. For example, hospitals that started in January 2015 or earlier had a full 9-month intervention period (since we required 3 months of follow-up for each patient, and our claims data extended through December 2015); those that started in April 2015 had only a 6-month intervention period. The mean intervention period was 7 months in our main analyses.

OUTCOMES

Our primary outcome was the change in standardized Medicare payments per episode. Secondary outcomes included changes in hospital case mix (based on the mean Chronic Conditions Warehouse score [a Medicare-provided comorbidity index that ranges from 0 to 27, with higher scores indicating more coexisting conditions]), per-hospital case volume, proportion of patients dually enrolled in Medicare and Medicaid (as an indicator of poverty), proportion of patients with disabling conditions, changes in the individual components of payment (for the index hospitalization, subsequent inpatient and outpatient care, and physician fees), and changes in length of stay, emergency department use, readmissions, and mortality.

STATISTICAL ANALYSIS

Hospital and market characteristics were compared between BPCI hospitals and matched control hospitals with the use of appropriate statistical tests. We used a difference-in-differences approach to examine changes in each outcome from the baseline period to the intervention period. To validate this approach we compared slopes for the 15-month period before the intervention between BPCI hospitals and control hospitals and found no significant differences in Medicare payments for all five conditions combined or any of the individual conditions (see Table S1 in the Supplementary Appendix, available with the full text of this article at NEJM.org). For each control hospital, the time periods were identical to those of the matched BPCI hospital. Analyses were run at the patient level, with each outcome in a separate model. Time was the primary predictor, coded as a binary variable for baseline versus intervention periods, along with BPCI status, and the interaction term between these indicators. The significance of the interaction term determined whether there had been a greater change in any outcome in patients at BPCI hospitals than in patients at control hospitals. Marginal models that use generalized estimating equations (the GENMOD procedure in the SAS statistical package) was the approach used to account for correlation among patients within hospitals. We included indicator variables for DRGs and used individual Chronic Conditions Warehouse scores to control for coexisting conditions.

For overall costs, we ran hospital-level analyses, since data on payments for outpatient services and physicians were available for only 20% of patients, whereas data on all other payments were available for 100% of patients. We used linear regression to estimate outpatient and physician payments for each hospital and separately to estimate mean "other" payments for each hospital. We then evaluated the resulting total adjusted mean costs in a linear regression model. Analyses were weighted by the number of patients in each hospital for each time period. A more detailed description is provided in the Supplementary Appendix.

All analyses were carried out first across all five conditions and then separately for each condition. We performed two sets of sensitivity analyses. First, we limited hospital-condition pairs to hospitals that did not drop out of BPCI early for the condition in question (287 hospitals). Second, to determine whether performance improved with the duration of participation, we limited the analytic sample to hospitals that joined the program by January 1, 2015 (45 hospitals), examining performance after a 6-month run-in period and using all available data through the end of 2015. For this group, our mean intervention period was 16 months.

Our primary outcome, the change in total Medicare payments per episode, was tested in six patient cohorts, so the Hochberg–Benjamini method of using a stepped P value to adjust for multiple comparisons[9] allowed us to determine statistical significance. In six comparisons (the five individual conditions plus all conditions combined), the largest change (rank 1) in P value was 0.05÷6, which equaled a corrected P value threshold of 0.008 (see the Supplementary Appendix). Secondary end points and analyses should be considered exploratory.

RESULTS

HOSPITAL AND MARKET CHARACTERISTICS

A total of 125 hospitals participated in the program for CHF, 105 hospitals for pneumonia, 101 hospitals for COPD, 88 hospitals for sepsis, and 73 hospitals for AMI. Of these hospitals, 62.6% participated in more than one of these bundles (Table 1, and Table S2 in the Supplementary Appendix). Overall, as compared with the sample of all eligible U.S. hospitals, BPCI hospitals were more likely to be nonprofit, urban, teaching hospitals with a larger number of beds. The matched control hospitals were similar to the BPCI hospitals with respect to each of these characteristics, as well as market characteristics and condition-specific, Medicare-allowed payments at baseline (Table 1, and Table S2 in the Supplementary Appendix).

CHANGES IN MEDICARE PAYMENTS

At baseline, the mean Medicare payment per episode across the five conditions for the BPCI hospitals was $24,280, which decreased to $23,993 during the intervention period (difference, −$286; P=0.41) (Fig. 1 and Table 2). (Differences may not sum precisely because of rounding.) Control hospitals had a mean Medicare payment of $23,901 per episode at baseline, which decreased to $23,503 during the intervention period (difference, −$398; P=0.08; difference in differences, $112; P=0.79). The findings were largely similar for exploratory analyses examining each of the individual conditions (Fig. 1, and Table S3 in the Supplementary Appendix). However, payments for pneumonia decreased by $1,495 for BPCI hospitals versus $278 for control hospitals (difference in differences, $1,216; P=0.03), which was not significant after adjustment for multiple comparisons.

CHANGES IN SECONDARY OUTCOMES

The individual components of Medicare payments per episode, including payments for the index hospitalization, readmissions, skilled nursing facilities, inpatient rehabilitation, home health care, and physician fees, did not differ significantly from baseline to the intervention period in the combined sample (Table 2). The findings were largely similar for exploratory analyses examining the individual conditions (Table S3 in the Supplementary Appendix).

In our combined analyses, patients in the BPCI hospitals and those in the control hospitals did not differ significantly with respect to changes from baseline to the intervention period in demographic characteristics or measures of medical or social risk. For example, in BPCI hospitals, the Chronic Conditions Warehouse score increased from 8.91 to 9.14, as compared with an increase from 8.94 to 9.10 in control hospitals (difference in differences, 0.07; P=0.17) (Table 3). Findings were similar for the individ-

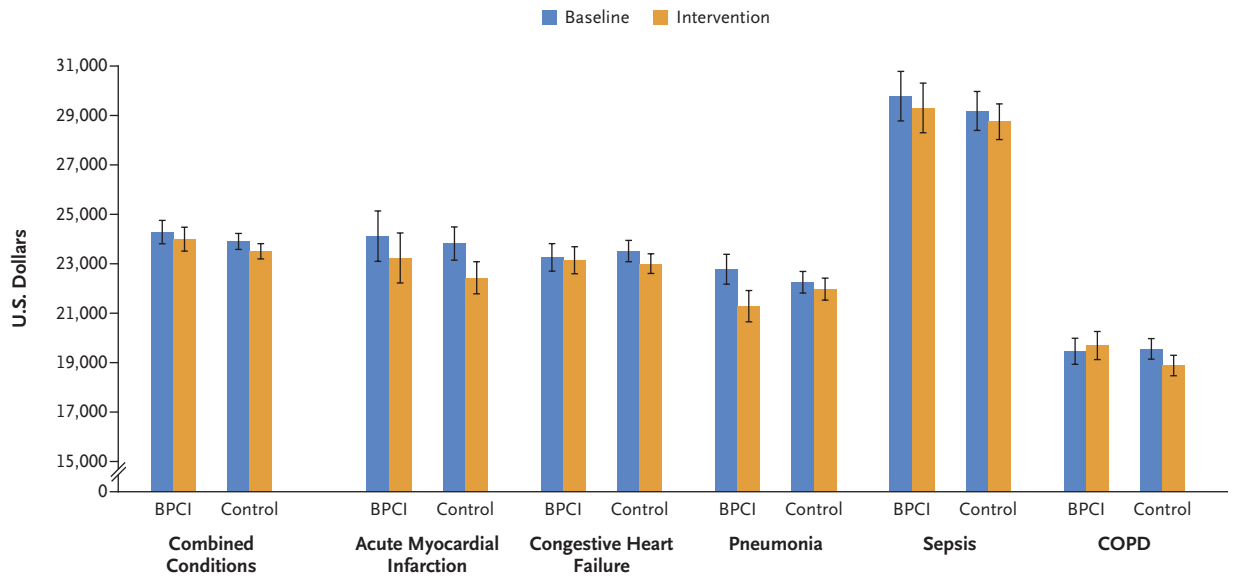

Figure 1. Standardized, Medicare-Allowed Payments for Hospitals Participating and Those Not Participating in the Bundled Payments for Care Improvement (BPCI) Initiative.

Shown are baseline and intervention Medicare payments to BPCI hospitals and matched control hospitals. Costs have been adjusted for patient-level coexisting conditions on the basis of Medicare's Chronic Conditions Warehouse data. COPD denotes chronic obstructive pulmonary disease. The I bars indicate standard errors.

ual conditions (Table S4 in the Supplementary Appendix).

There were no significant overall differences between BPCI hospitals and control hospitals in the change from baseline to the intervention period for length of stay, emergency department visits or readmissions within 30 or 90 days after discharge, or mortality within 30 or 90 days after admission (Table 4). For example, although both BPCI and control hospitals had increases in 90-day mortality during the study period, the degree of increase did not differ significantly between the two types of hospitals. In BPCI hospitals, the increase was from 15.1 to 17.8% (difference, 2.8 percentage points; P<0.001), and in control hospitals, the increase was from 14.5 to 17.0% (difference, 2.4 percentage points; P<0.001); the difference in differences was 0.3 percentage points (P=0.52). The findings were largely similar in exploratory analyses examining the individual conditions. However, in the AMI cohort, 90-day mortality decreased in BPCI hospitals relative to control hospitals, and in the COPD cohort, mortality at 90 days and the readmission rate at 30 days increased in BPCI hospitals relative to control hospitals (Table S5 in the Supplementary Appendix).

SENSITIVITY ANALYSES

When we limited our sample to hospitals that did not drop out of BPCI, the findings were qualitatively similar, with the exception of a drop in quarterly volume for BPCI hospitals as compared with control hospitals (Tables S6, S7, and S8 in the Supplementary Appendix). Finally, to allow time for hospitals to incorporate care redesign, we excluded the first 6 months of the intervention period and examined changes thereafter. Again, the results were qualitatively similar, aside from an isolated differential decrease in spending on inpatient rehabilitation for BPCI hospitals (–$135) as compared with control hospitals ($40); difference in differences, –$175; P=0.002 (Tables S9, S10, and S11 in the Supplementary Appendix).

DISCUSSION

We found that hospital participation in episode-based payment for the five most commonly selected medical conditions under BPCI, account-

Table 2. Changes in Medicare Payments per Episode.*

Variable	BPCI Hospitals				Matched Control Hospitals			Comparison	
	Baseline	Intervention Period	Difference	P Value	Baseline	Intervention Period	Difference	P Value	Difference in Differences	P Value
	U.S. dollars				*U.S. dollars*				*U.S. dollars*	
Total 90-day Medicare payments	24,280	23,993	−286	0.41	23,901	23,503	−398	0.08	112	0.79
Index hospitalization	7,687	7,696	9	0.94	7,576	7,566	−9	0.90	18	0.89
Readmission	5,449	5,226	−223	0.07	5,510	5,122	−388	<0.001	165	0.26
Skilled nursing	5,391	5,392	1	0.99	5,167	5,123	−44	0.62	45	0.78
Inpatient rehabilitation	406	379	−27	0.61	420	433	14	0.70	−40	0.52
Long-term care hospital	639	430	−209	0.01	551	445	−106	0.06	−103	0.30
Home health agency	1,193	1,068	−124	<0.001	1,212	1,111	−101	<0.001	−24	0.47
Physician fees	2,783	2,951	167	0.004	2,720	2,838	119	0.002	49	0.49
Nonphysician outpatient fees	732	851	120	<0.001	746	864	118	<0.001	2	0.96

* Episodes included hospitalization plus 90 days after discharge. Costs have been adjusted with the use of data on patient-level coexisting conditions from the Medicare Chronic Conditions Warehouse. For similar data on each condition individually, see Table S3 in the Supplementary Appendix. The differences may not sum precisely because of rounding.

ing for nearly two thirds of patients enrolled in medical bundles in this program, was not associated with significant changes in total or component Medicare episode payments, clinical complexity, length of stay, emergency department use, readmissions, or mortality.

Our findings differ from the prior study of BPCI for total joint replacement, which showed that participating hospitals were successful in lowering overall Medicare payments.[4] Medical bundles might differ from joint replacement in a few ways: joint replacement is elective, and patients tend to be younger (30% of patients were older than 80 years of age in the prior study of total joint replacement, whereas 55% of patients were over the age of 80 years in our cohort) and have lower rates of poverty and disability (12.7% of patients were dually enrolled and 10.5% were disabled in the study of joint replacement, whereas 25.1% were dually enrolled and 21.3% were disabled in our study).[4] As a result of these complexities, patients admitted for medical conditions may have had post–acute care needs that were less amenable to intervention.[10-12] However, in analyses performed by the Lewin Group under a federal contract, payments for joint replacements associated with fracture — presumably a nonelective procedure in a medically more complex subgroup of patients undergoing joint replacement — were also lower for BPCI hospitals, suggesting that patient characteristics alone may not account for our findings.[5]

Medicare payments for the inpatient stay for any particular diagnosis are largely constant across hospitals under the DRG system. Consequently, most savings under BPCI would need to come from post–acute care services.[13,14] Therefore, another possibility for the failure of BPCI hospitals to reduce allowed payments is a lack of ability to influence care provided by skilled nursing facilities, inpatient rehabilitation facilities, long-term care hospitals, or home health agencies. In the absence of relationships with these providers, hospitals may have little say in what happens to patients once they enter a post–acute care setting. However, recent research suggests that partnerships between hospitals and post–acute care providers are becoming more common.[15]

There have been prior studies of bundling, though few for medical conditions. In 1991, CMS introduced a 5-year demonstration project for coronary-artery bypass grafting in seven hos-

Table 3. Changes in Volume and Case Mix.*

Variable	BPCI Hospitals				Matched Control Hospitals				Comparison	
	Baseline	Intervention Period	Difference†	P Value	Baseline	Intervention Period	Difference†	P Value	Difference in Differences	P Value
No. of episodes	39,442	38,839			70,780	76,167				
Mean no. of discharges per hospital per quarter	65.4	62.1	−3.3	0.29	72.6	76.0	3.4	0.13	−6.7	0.08
Patient characteristics (%)										
≤64 yr of age	8.7	9.4	0.8	0.05	9.7	9.6	−0.1	0.71	0.9	0.07
65–79 yr of age	35.8	35.3	−0.5	0.46	35.7	35.4	−0.3	0.41	−0.1	0.84
≥80 yr of age	55.5	55.2	−0.3	0.68	54.6	55.1	0.4	0.33	−0.7	0.40
Female sex	55.1	55.0	−0.1	0.84	53.9	54.0	0.1	0.77	−0.2	0.74
Medicaid beneficiary	25.1	25.2	0.1	0.88	24.5	24.6	0.1	0.86	0.0	0.98
Disabled without ESRD	21.3	22.3	1.0	0.10	23.1	22.8	−0.3	0.46	1.3	0.07
Race or ethnic group‡										
White	83.9	83.1	−0.8	0.25	83.6	83.3	−0.3	0.49	−0.4	0.61
Black	10.7	11.1	0.4	0.49	10.9	11.2	0.3	0.48	0.1	0.88
Hispanic	2.3	2.3	0.0	0.95	1.9	1.8	−0.1	0.42	0.1	0.67
Unknown or other	3.1	3.5	0.4	0.11	3.6	3.7	0.1	0.46	0.2	0.44
Mean Chronic Conditions Warehouse score§	8.91	9.14	0.22	<0.001	8.94	9.10	0.16	<0.001	0.07	0.17
Clinical complexity (%)										
DRG with MCC	39.5	41.2	1.7	0.02	40.8	42.8	2.0	<0.001	−0.4	0.71
DRG with CC	44.8	43.4	−1.3	0.05	43.5	42.3	−1.2	0.01	−0.1	0.91
DRG without CC	15.7	15.4	−0.3	0.46	15.7	14.9	−0.8	0.02	0.5	0.41

* CC denotes complication or coexisting condition, DRG diagnosis-related group, ESRD end-stage renal disease, and MCC major complication or coexisting condition. For similar data on each condition individually, see Table S4 in the Supplementary Appendix. The differences may not sum precisely because of rounding. Percentages may not sum to 100 because of rounding.

† Values for the difference between percentages are percentage points.

‡ Race or ethnic group was determined on the basis of Medicare enrollment data.

§ The Chronic Conditions Warehouse score is a Medicare-supplied comorbidity measure that ranges from 0 to 27, with higher scores indicating more coexisting conditions.

Table 4. Changes in Length of Stay, Emergency Department (ED) Use, Hospital Readmissions, and Mortality.*

Variable	BPCI Hospitals				Matched Control Hospitals				Comparison	
	Baseline	Intervention Period	Difference†	P Value	Baseline	Intervention Period	Difference†	P Value	Difference in Differences	P Value
Mean length of stay (days)	5.6	5.6	−0.1	0.41	5.6	5.5	−0.1	0.13	0.0	0.92
ED use (%)										
At 30 days	9.9	10.2	0.3	0.68	10.9	10.2	−0.7	0.14	1.0	0.23
At 90 days	19.5	17.9	−1.6	0.10	20.9	18.4	−2.5	<0.001	0.9	0.44
Readmission (%)										
At 30 days	16.8	16.3	−0.5	0.17	17.1	16.7	−0.4	0.22	−0.1	0.79
At 90 days	32.4	30.7	−1.7	<0.001	32.5	30.8	−1.7	<0.001	0.0	0.97
Mortality (%)										
At 30 days	7.8	8.5	0.6	0.01	7.4	8.2	0.8	<0.001	−0.2	0.64
At 90 days	15.1	17.8	2.8	<0.001	14.5	17.0	2.4	<0.001	0.3	0.52

* Outcomes have been adjusted with the use of data on patient-level coexisting conditions from the Medicare Chronic Conditions Warehouse. For similar data on each condition individually, see Table S5 in the Supplementary Appendix. The differences may not sum precisely because of rounding.
† Values for the difference between percentages are percentage points.

pitals, which showed cost savings of about 10%, with reduced mortality and complication rates. In 2009, CMS launched the Acute Care Episode (ACE) demonstration project for cardiac and orthopedic care, which resulted in some savings for both types of care.[16] More recently, national and single-center studies have shown savings for total joint replacement under BPCI,[4,17,18] a single-center study showed no savings in a COPD bundle under BPCI,[19] and three annual federal evaluations of BPCI reported qualitatively similar findings to ours for medical conditions.[5] Differences between the federal findings and ours are probably related to slight differences in baseline and follow-up periods, variables used in matching, selection of control hospitals, and availability of outpatient data.

There are limitations to our findings. First, BPCI is a voluntary program, and we examined only five conditions, so generalizability to other conditions or mandatory models is uncertain. Second, we focused on hospital participants; though they account for the majority of model 2 participants,[1] patterns may differ for physician practices that participate in the program. Third, our lists of hospital participants were obtained from CMMI and were not verified by the hospitals themselves, which may have introduced error. Fourth, since CMMI has not released data on target pricing or on hospitals' savings or losses under the program, we could evaluate only the effect of BPCI on patients and their outcomes, not its effect on hospital finances. We recommend that interested parties view the federal reports[5,6] for aggregate findings on hospital finances, as well as additional information gleaned from federally administered surveys of participants and patients.

Fifth, there are no validated risk-adjustment models for 90-day outcomes, so we adapted 30-day models for that purpose. Sixth, we used a relatively short baseline period; during this period, hospitals were surely aware that they were nearing the beginning of the intervention and may have been preparing to redesign care. However, our finding that baseline trends did not differ between intervention and control hospitals is reassuring. Also, since hospitals are judged against their own historical baseline, there was no incentive to begin reducing costs before the start date. Finally, we had a limited follow-up period, and longer-term follow-up may be neces-

sary to fully evaluate how care evolves.[20] Nevertheless, similar follow-up with a similar sample size was sufficient for the prior study to identify an effect of BPCI on joint replacement.[4]

In summary, hospital participation in five common medical bundles under BPCI, as compared with nonparticipation, was not associated with changes from baseline in total Medicare payments per episode, case complexity, length of stay, emergency department use, hospital readmission, or mortality. Bundling of services to encourage more efficient care has great face validity and enjoys bipartisan support. For such bundling to work for medical conditions, however, more time, new care strategies and partnerships, or additional incentives may be required.

Supported by the Commonwealth Fund.

Disclosure forms provided by the authors are available with the full text of this article at NEJM.org.

We thank Emily Crawford, M.A., for her assistance with data collection and Laura Dummit, M.S.P.H., and Grecia Marrufo, Ph.D., for their comments on an earlier version of the manuscript.

REFERENCES

1. Bundled Payments for Care Improvement (BPCI) initiative: general information. Baltimore: Centers for Medicare and Medicaid Services, 2017 (http://innovation.cms.gov/initiatives/bundled-payments/).
2. Mechanic RW. Medicare's bundled payment initiatives: considerations for providers. Washington, DC: American Hospital Association, 2016 (https://www.aha.org/system/files/content/16/issbrief-bundledpmt.pdf).
3. BPCI Advanced home page. Baltimore: Centers for Medicare and Medicaid Services, 2017 (https://innovation.cms.gov/initiatives/bpci-advanced).
4. Dummit LA, Kahvecioglu D, Marrufo G, et al. Association between hospital participation in a Medicare bundled payment initiative and payments and quality outcomes for lower extremity joint replacement episodes. JAMA 2016;316:1267-78.
5. Dummit L, Marrufo G, Marshall J, et al. CMS Bundled Payments for Care Improvement initiative models 2-4: year 3 evaluation and monitoring annual report. Falls Church, VA: Lewin Group, October 2017 (https://downloads.cms.gov/files/cmmi/bpci-models2-4yr3evalrpt.pdf).
6. Dummit L, Marrufo G, Marshall J, et al. CMS Bundled Payments for Care Improvement initiative models 2-4: year 2 evaluation and monitoring annual report. Falls Church, VA: Lewin Group, August 2016 (https://innovation.cms.gov/Files/reports/bpci-models2-4-yr2evalrpt.pdf).
7. Herfindahl–Hirschman Index. Washington, DC: Department of Justice, July 29, 2015 (https://www.justice.gov/atr/herfindahl-hirschman-index).
8. Austin PC. Optimal caliper widths for propensity-score matching when estimating differences in means and differences in proportions in observational studies. Pharm Stat 2011;10:150-61.
9. Hochberg Y, Benjamini Y. More powerful procedures for multiple significance testing. Stat Med 1990;9:811-8.
10. Samson LW, Chen LM, Epstein AM, Joynt Maddox KE. Dually enrolled beneficiaries have higher episode costs on the Medicare Spending Per Beneficiary measure. Health Aff (Millwood) 2018;37:86-94.
11. Das A, Norton EC, Miller DC, Chen LM. Association of postdischarge spending and performance on new episode-based spending measure. JAMA Intern Med 2016;176:117-9.
12. Kasper J, O'Malley Watts M, Lyons B. Chronic disease and co-morbidity among dual eligibles: implications for patterns of Medicaid and Medicare service use and spending. Washington, DC: Kaiser Commission on Medicaid and the Uninsured, Henry J. Kaiser Family Foundation, July 2010 (https://kaiserfamilyfoundation.files.wordpress.com/2013/01/8081.pdf).
13. Tsai TC, Joynt KE, Wild RC, Orav EJ, Jha AK. Medicare's Bundled Payment initiative: most hospitals are focused on a few high-volume conditions. Health Aff (Millwood) 2015;34:371-80.
14. Institute of Medicine. Interim Report of the Committee on Geographic Variation in Health Care Spending and Promotion of High-Value Care: preliminary committee observations. Washington, DC: National Academies Press, 2013.
15. McHugh JP, Foster A, Mor V, et al. Reducing hospital readmissions through preferred networks of skilled nursing facilities. Health Aff (Millwood) 2017;36:1591-8.
16. Evaluation of the Medicare Acute Care Episode (ACE) demonstration. Baltimore: Centers for Medicare & Medicaid Services, May 31, 2013 (http://downloads.cms.gov/files/cmmi/ACE-EvaluationReport-Final-5-2-14.pdf).
17. Navathe AS, Troxel AB, Liao JM, et al. Cost of joint replacement using bundled payment models. JAMA Intern Med 2017;177:214-22.
18. Iorio R, Clair AJ, Inneh IA, Slover JD, Bosco JA, Zuckerman JD. Early results of Medicare's bundled payment initiative for a 90-day total joint arthroplasty episode of care. J Arthroplasty 2016;31:343-50.
19. Bhatt SP, Wells JM, Iyer AS, et al. Results of a Medicare Bundled Payments for Care Improvement initiative for chronic obstructive pulmonary disease readmissions. Ann Am Thorac Soc 2017;14:643-8.
20. Dundon JM, Bosco J, Slover J, Yu S, Sayeed Y, Iorio R. Improvement in total joint replacement quality metrics: year one versus year three of the Bundled Payments for Care Improvement Initiative. J Bone Joint Surg Am 2016;98:1949-53.