

Original Article

Medicare's Bundled Payment Initiatives for Hospital-Initiated Episodes: Evidence and Evolution

CHRISTINE A. YEE,^{*,†,‡} STEVEN D. PIZER,^{*,‡}
and AUSTIN FRAKT^{*,‡,§}

**Partnered Evidence-based Policy Resource Center, VA Boston Healthcare System; †University of Maryland Baltimore County; ‡School of Public Health, Boston University; §T.H. Chan School of Public Health, Harvard University*

Policy Points:

- Evidence suggests that bundled payment contracting can slow the growth of payer costs relative to fee-for-service contracting, although bundled payment models may not reduce absolute costs.
- Bundled payments may be more effective than fee-for-service payments in containing costs for certain medical conditions.
- For the most part, Medicare's bundled payment initiatives have not been associated with a worsening of quality in terms of readmissions, emergency department use, and mortality. Some evidence suggests a worsening of other quality measures for certain medical conditions.
- Bundled payment contracting involves trade-offs: Expanding a bundle's scope and duration may better contain costs, but a more comprehensive bundle may be less attractive to providers, reducing their willingness to accept it as an alternative to fee-for-service payment.

Context: Bundled payments have been promoted as an alternative to fee-for-service payments that can mitigate the incentives for service volume under the fee-for-service model. As Medicare has gained experience with bundled payments, it has widened their scope and increased their duration. However, there have been few reviews of the empirical literature on the impact of Medicare's bundled payment programs on cost, resource use, utilization, and quality.

Methods: We examined the history and features of 16 of Medicare's bundled payment programs involving hospital-initiated episodes of care and conducted

a literature review of articles about those programs. Database and additional searches yielded 1,479 articles. We evaluate the studies' methodological quality and summarize the quantitative findings about Medicare expenditures and quality of care from 37 studies that used higher-quality research designs.

Findings: Medicare's bundled payment initiatives have varied in their design features, such as episode scope and duration. Many initiatives were associated with little to no reduction in Medicare expenditures, unless large pricing discounts for providers were negotiated in advance. Initiatives that included post-acute care services were associated with lower expenditures for certain conditions. Hospitals may have been able to reduce internal production costs with help from physicians via gainsharing. Most initiatives were not associated with significant changes in quality of care, as measured by readmission and mortality rates. Of the significant changes in readmission rates, the results were mixed, showing increases and decreases associated with bundled payments. Some evidence suggested that worse patient outcomes were associated bundled payments, although most results were not statistically significant. Results on case-mix selection were mixed: Several initiatives were associated with reductions in episode severity, whereas others were associated with little change.

Conclusions: Bundled payments for hospital-initiated episodes may be a good alternative to fee-for-service payments. Bundled payments can help slow the growth of payer spending, although they do not necessarily reduce absolute spending. They are associated with lower provider production costs, and there is no overwhelming evidence of compromised quality. However, designing a bundled payment contract that is attractive to both providers and payers proves to be a challenge.

Keywords: Bundled payment, prospective payment, alternative payment model, cost containment, Medicare.

TO CONTAIN THE GROWTH OF HEALTH CARE SPENDING, insurers and providers are experimenting with alternative payment structures that mitigate the incentive to increase volume that is inherent in traditional fee-for-service payment. One such alternative is bundled payment contracting, in which the payer pays a fixed amount for services associated with an episode of care. There is an expectation that bundled payments will lead to lower spending without affecting quality.¹⁻⁵

Medicare has a long history of experience with bundled payments (more so than private payers and employers), dating back to 1983. Interest in bundled payment models has increased since Section 3023 of the Patient Protection and Affordable Care Act (ACA), as modified by

Section 10308, required the Secretary of Health and Human Services to test and evaluate the use of bundled payments for hospital-initiated episodes. Under this authority, deliberations about continuing, cutting, or modifying bundled payment initiatives are ongoing within the Medicare program. Each bundled payment initiative has had a distinctive set of features that potentially affects the expected results of bundled payments. Because the Centers for Medicare and Medicaid Services (CMS) have experimented by choosing different features, it is important to evaluate which features are associated with specific outcomes.

The focus of this review is Medicare's experience with bundled payment initiatives for episodes of care initiated by hospital inpatient admissions. We focus on hospital-initiated episodes because those types of episodes capture a large share of Medicare spending and Medicare has experimented the most with bundled payments for them. Based on our review of the literature, we note several key findings:

- Bundled payment initiatives have reduced Medicare expenditures for certain medical conditions, but in many cases, spending has not been reduced. The magnitude depends on how much of a discount is applied to status quo fees and whether the discounted episode price is known by providers prior to rendering services. In more recent initiatives, when there were savings, they were often achieved in post-acute care spending.
- The literature indicates that hospitals' internal production costs have decreased under many bundled payment initiatives. This is important because it feeds into hospitals' decisions to participate in bundled payment contracting—which are often, but not always, voluntary—and it suggests potential future reductions in payer expenditures.
- In general, bundled payments have not altered (consistently in one direction) readmission rates, mortality rates, or emergency department (ED) use. However, there are some cases for which functioning has worsened.
- Evidence on potential unintended consequences indicate that some providers may be more selective in the patients they treat under bundled payments—shifting the composition of patients toward lower-cost patients. However, high-quality evidence on case mix severity is limited. It is even more limited for episode volume.

- Medicare has gained experience from its bundled payment programs and, with subsequent ones, it has adjusted episode definitions, lengthening their duration and widening their scope.

This review enhances the body of knowledge found in prior reviews of bundled payment studies¹⁻⁵ in several ways. First, we assessed how recent initiatives promulgated by the ACA, such as the Bundled Payments for Care Improvement (BPCI) and Comprehensive Care for Joint Replacement (CJR) initiatives, compare to older initiatives. Second, through our exploration of the variation in the features of bundled payment initiatives, we investigated which features have been associated with changes in spending and quality. Third, we provided side-by-side comparisons of the estimated effects of Medicare's bundled payment initiatives (including 95% confidence intervals when applicable), summarizing key findings on spending, resource use, utilization, and quality. Finally, we focused on studies with higher quality research designs. Previously published reviews of bundled payments included studies with research designs that do not have a control group; these can yield larger and possibly biased estimates of the impact of bundled payments. In contrast, we have limited our review to studies that assess bundled payment implementation relative to a plausible counterfactual or control group. Before discussing the literature review, let us first consider what bundled payments are, their history within Medicare, and the features of Medicare's various bundled payment initiatives.

What Are Bundled Payments?

For the purposes of this review, *bundled payment contract* means the payer pays one lump-sum price for multiple medical services that are provided to a patient during a predefined episode of care. The *episode of care* typically begins with a patient's *index event* and continues for a specified span of time. For example, the index event could be a hospital admission for a hip replacement and the span of time could be 90 days after discharge. Sometimes, the episode can include a short period before the index event, such as 14 days prior to hospitalization, as in the Hospital Gainsharing (HG) and Physician Hospital Collaboration (PHC) initiatives. Sometimes, as in the Bundled Payments for Care Improvement (BPCI) Model 3, the episode of care may not include the index event and may only include the post-acute care period. In more recent

programs (ie, many of those since 2008), time spans of 30, 60, or 90 days postdischarge have been typical episode durations.

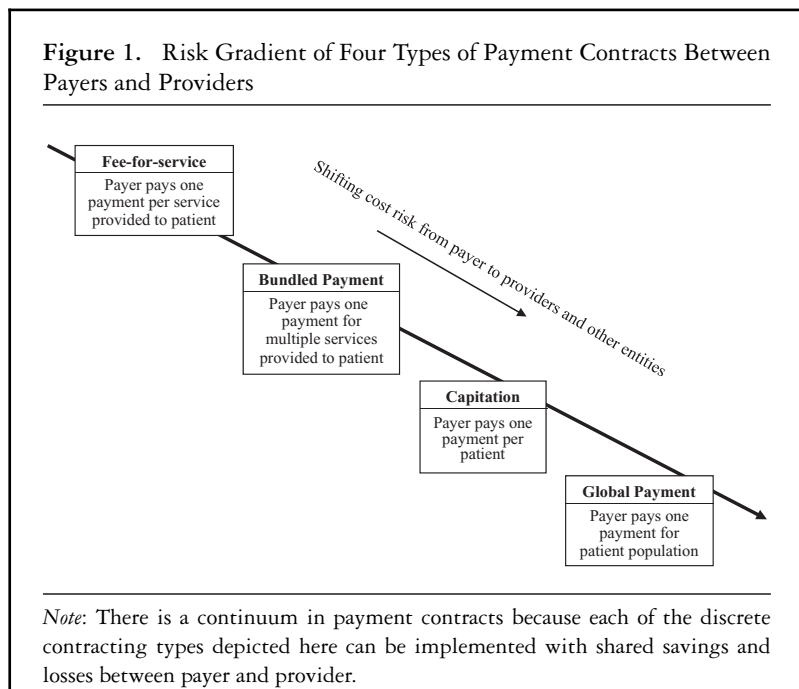
The *bundle* defines the types of services during the episode of care that are to be compensated with a single, comprehensive payment. Depending on the contract, the bundle can include (but is not limited to) hospital, physician, and/or post-acute care services.

Relative to the traditional fee-for-service arrangement, a bundled payment puts more financial risk on the contract *awardee*, which is the provider or collection of providers responsible financially for gains and losses. However, these at-risk awardees do not bear the financial risk for all the care that could be related to a patient's condition. The payer still pays for services that are outside of the bundle definition or provided after the episode of care. These services may be related to severity that is not accounted for in the bundle payment. Additionally, the payer is at risk for the number of episodes of care in a population. In fact, bundled payments exist within a continuum of risk-sharing mechanisms as depicted in Figure 1 (described later).

Why Bundled Payments?

Relative to comparable countries, the United States has higher medical spending per capita with largely inferior health outcomes, access, and efficiency.^{6,7} Slowing health care spending growth and improving care quality have become national priorities, prompting health care payers such as Medicare and private insurers to experiment with new ways of paying health care providers.

Bundled payments exist within a continuum of risk-sharing mechanisms (Figure 1) and mitigate some of the adverse characteristics of the payment models at the extremes of that continuum: fee-for-service and global payment contracting. With fee-for-service contracting, payers pay providers an amount for each service provided. Fee-for-service contracting encourages a greater volume of services and does not hold the treating provider responsible for potentially avoidable services. With global payment contracting, payers pay providers an amount to take care of a certain population, regardless of the type or quantity of services provided. Global payments put significant financial risk on providers, potentially encouraging them to provide too few services (as this would increase profits).



Another form of contracting is capitation, in which payers pay providers a set amount per patient regardless of services rendered. Relative to global payments, capitation shifts more financial risk to payers because payers bear the financial cost when the number of patients seeking treatment increases. Bundled payment, which shifts more financial risk to payers than capitation and global payments, has been employed by Medicare as a middle ground that balances payer and provider financial risk.⁸

History of Medicare's Bundled Payment Initiatives

Payment for hospitalizations in the United States has been a focus of experimentation for at least four decades. In 1980 (15 years into the Medicare program), hospital payments were 71% of all Medicare payments and were increasing at an annual rate of 15.4%.⁹ This was a time when

physicians were not yet subject to a fee schedule and hospitals were paid based on a retrospective cost-plus reimbursement system. Altman and Eichenholz¹⁰ characterized this as a “blank check” environment that led to an inefficient use of resources.

In 1982, Congress passed the Tax Equity and Fiscal Responsibility Act, which altered the way hospitals were paid by Medicare. The Inpatient Prospective Payment System (IPPS) was introduced a year later. IPPS was Medicare’s first hospital bundled payment program. It bundles the payment for hospital services associated with an eligible inpatient admission and prospectively sets the amount for each type of diagnosis. This payment structure incentivizes hospitals to reduce production costs and lengths of stay have declined since its implementation.^{11,12}

After launching IPPS, Medicare implemented smaller-scale bundled payment initiatives, such as the Medicare Participating Heart Bypass Center (HBC) demonstration in 1991, the Medicare Cataract Surgery Alternate Payment (CSAP) demonstration in 1993, the Hospital Gain-sharing (HG) demonstration in 2008, the Physician Hospital Collaboration (PHC) demonstration in 2009, and the Acute Care Episode (ACE) demonstration in 2009. In various ways, these initiatives expanded the scope of the bundle and the length of the episode of care beyond the inpatient stay.

The ACA created the Center for Medicare and Medicaid Innovation (CMMI) within CMS. CMMI’s mission has been to develop, implement, and test various ways of paying providers. Such initiatives include BPCI Models 1-4 and the Comprehensive Care for Joint Replacement (CJR) initiative, which share properties of earlier initiatives.

Many bundled payment initiatives were designed to incentivize hospitals to reduce production costs. Hospitals did so in several ways, such as negotiating lower prices for supplies and devices, reducing the number of nurse-hours dedicated to a case, and encouraging physicians to make medical decisions that are most cost effective for the hospital.¹³⁻¹⁵ In some initiatives, hospital managers were legally allowed to incentivize physicians with additional compensation tied to the amount by which hospital production costs were reduced. This ability to gainshare was one of the reasons that hospitals volunteered to participate in bundled payment contracts, even offering or agreeing to discounts in Medicare reimbursement. With expectations of decreasing internal production costs, hospitals were willing to accept lower payments per episode of care.

Traditionally, US physicians and hospitals have operated independently, meaning physicians and hospitals have separate payment contracts with payers and their own costs of production. However, the quantity of services provided by hospitals often is determined by physicians, who, traditionally, have not borne the cost consequences associated with those services. In contrast, bundled payments with gainsharing can align physician incentives with the business practices of hospitals. Under several demonstrations (eg, ACE), this alignment was so important to hospitals that they bore all financial risk and offered physicians gain-sharing bonuses to secure physician participation.^{14,15}

One study reported skepticism among physicians about gainsharing: they expressed concerns about whether they would actually be paid more, whether more work would be required to secure bonuses, and whether their decisions would be closely scrutinized.¹⁶ Physicians also were concerned about the potential for malpractice lawsuits if patients learned of their gainsharing and suspected it influenced medical decision-making. However, bundled payments gained more traction among physicians as they learned that early physician adopters had earned bonuses.¹⁶ In addition, as part of several bundled payment initiatives, Medicare promised to advertise and promote participating providers to patients. In the ACE demonstration, for example, Medicare encouraged patients to go to participating providers by giving the patients up to 50% of Medicare's cost savings.¹⁷

Almost all bundled payment initiatives launched after IPPS have included physician services in the bundle definition. Relative to gainsharing, this approach more directly involves physicians in the financial risk of care services. In more recent initiatives, contractual relationships have gone beyond physicians and hospitals. HG, PHC, BPCI Models 2 and 3, CJR, and BPCI Advanced have included post-acute care services in the bundle definition. These incentives can encourage post-acute care facilities to coordinate with physicians and hospitals, creating networks and pathways to coordinate care.

Features of Medicare's Bundled Payment Initiatives

Since IPPS was initiated in 1983, there have been 16 bundled payment initiatives centered around hospital admissions: two in the early 1990s,

three just before the ACA, nine initiated under the authority of the ACA, and one initiated in 2018. Five of the nine ACA initiatives are currently ongoing. Four others—the Acute Myocardial Infarction (AMI) Model; the Coronary Artery Bypass Graft (CABG) Model; the Surgical Hip and Femur Fracture Treatment (SHFFT) Model; and the Cardiac Rehabilitation (CR) Incentive Payment Model—were canceled by the Trump administration. Although we have not reported findings on the impact of the canceled initiatives, we have included them in our discussion of initiative features to provide context, and because they (or initiatives of similar structure) might be revisited in the future.

Table 1 describes eight features of the 16 initiatives: 1) whether the initiative was a voluntary program, 2) the expected Medicare cost savings for Part A and B services assuming no change in utilization (ie, the discount off status quo reimbursement rates), 3) the medical conditions included in the initiative, 4) the types of provider services covered by the bundle, 5) the duration of the episode, 6) when payment amounts were determined and administered, 7) the location of participants, and 8) the number of participants as of February 2017 and May 2019. These key features of the 16 initiatives, in combination, can alter the incentive structure for providers and may influence the effectiveness of an initiative and the selection of participants. The features of the ACA-promulgated initiatives that were not canceled resemble those of earlier initiatives in many ways: BPCI Advanced, CJR, and BPCI Model 2 features are similar in design to the HG and PHC demonstrations. BPCI Model 4 was similar in design to HBC and somewhat similar to ACE; and BPCI Model 1 was a slight deviation from IPPS. We discuss the features captured in Table 1 in the following sections.

Voluntary Initiatives

With the exceptions of IPPS, CJR, and the four canceled ACA initiatives, bundled payment initiatives have been voluntary (see Table 1). This means that Medicare has solicited providers to participate (via a proposal process) and often selected a subset of the provider applicants to participate in a program.

CJR is not voluntary. CMS currently requires certain hospitals in 34 metropolitan areas to participate in this program and allows voluntary participation in 33 other areas. The four canceled initiatives were similar

Initiative Name	Time Frame ^a	Voluntary	Discount off Medicare Part A & B Fees ^b	Conditions	Services in Bundle ^c	Episode Duration ^d	Payment ^e	Geographic Area	Hospital Participants 2017 ^f	Hospital Participants 2019 ^g
Inpatient Prospective Payment System (IPPS)	1991-1996	Y	A+B: 9.7%-36.7%	CABG (DRG 106, 107)	H, PS, PG, O, R	72 hours ^h	P	GA, MI, OH, MA, TX, OR, IN	7	7
Medicare Participating Heart Bypass Center Demonstration (HBC)	1993-1996	Y	A: 0% B: 2%-5%	Outpatient cataract surgery	H, PS, PG, O	120 days + pre-op	P	OH, TX, AZ	4	4
Alternate Payment Demonstration (CSAP)	2008-2011	Y	BN	Up to participant	H, PA, PG, PAC, O, R	30 days + 14 days pre-op	P	NY, WV	2	2
Medicare Hospital Gainsharing Demonstration (HG)	2009-2013	Y	A: 0.9%-8.25% B: 0%-3.3%	Cardiac (29 DRGs), orthopedic hip, knee (9 DRGs)	H, PA, PG, O	n/a	P	OK, TX, NM, CO	5	5

Continued

Table 1. *Continued*

Initiative Name	Time Frame ^a	Voluntary	Discount off Medicare Part A & B Fees ^b	Conditions	Services in Bundle ^c	Episode Duration ^d	Payment ^e	Geographic Area	Hospital Participants 2017 ^f	Hospital Participants 2019
Physician Hospital Collaboration Demonstration (PHC)	2009-2013	Y	BN	All MS-DRGs	H, PA, PG, PAC, O, R	90 days + 14 days pre-op	P	NJ	12	12
BPCI Model 1 (BPCIIm1)	2013-2016	Y	A: 0%-1% B: 0%	All MS-DRGs	H, PG	n/a	P	NJ, KS	24	13
BPCI Model 2 (BPCIIm2)	2013-2018	Y	A+B: 2%-3%	Up to 48 MS-DRGs	H, PA, PG, PAC, O, R	30, 60, 90 days	R	National	539	402
BPCI Model 3 (BPCIIm3)	2013-2018	Y	A+B: 2%-3%	Up to 48 MS-DRGs	H, PA, PG, PAC, O, R	30, 60, 90 days (excludes admission)	R	National	758	577
BPCI Model 4 (BPCIIm4)	2013-2018	Y	A+B: 2%-3%	Up to 48 MS-DRGs	H, PA, PG, O, R	30 days ^j	P	National	5	2
Comprehensive Care for Joint Replacement (CJR) Model Initiative	2016-2021	N	A+B: 2%-3%	Hip, knee (MS-DRG 469, 470)	H, PA, PG, PAC, O, R	90 days	R+Q	67 MSAs ^k	798	488

Continued

Table 1. Continued

Initiative Name	Time Frame ^a	Voluntary	Discount off Medicare Part A & B Fees ^b	Conditions	Services in Bundle ^c	Episode Duration ^d	Payment ^e	Geographic Area	Hospital Participants 2017 ^f	Hospital Participants 2019
BPCI Advanced	2018-2023	Y	A+B: 3%	Up to 33 MS-DRGs and 4 HCPCS	H, PA, PG, PAC, O, R	90 days	R+Q	National	0	1,281
Acute Myocardial Infarction (AMI) Model	Originally 2017–2021 ¹	N	n/a	AMI, AMI+PCI (MS-DRG 280–282, 246–251)	H, PA, PG, PAC, O, R	90 days	R+Q	98 MSAs	1,095	0
Coronary Artery Bypass Graft (CABG) Model	Originally 2017–2021 ¹	N	n/a	CABG (MS-DRG 231–236)	H, PA, PG, PAC, O, R	90 days	R+Q	98 MSAs	1,096	0
Surgical Hip and Femur Fracture Treatment (SHFFT) Model	Originally 2017–2021 ¹	N	n/a	Hip, femur fracture (MS-DRG 480–482)	H, PA, PG, PAC, O, R	90 days	R+Q	67 MSAs	865	0
Cardiac Rehabilitation (CR) Incentive Payment Model	Originally 2017–2021 ¹	N	n/a	Cardiac rehab & support services for AMI, CABG	H, PG, O	90 days	P ^m	90 MSAs	0	0

Continued

Table 1. *Continued*

The information in the table was provided by the Center for Medicare and Medicaid Innovation (CMMI). Abbreviations: *BPCI*, Bundled Payments for Care Improvement; *MS-DRG*, Medicare severity diagnosis related group; *CABG*, Coronary artery bypass grafting; *PCI*, Percutaneous coronary intervention; *MSA*, metropolitan statistical area.

^a*Time Frame* includes the first cohort through the last cohort.

^b*BN* = Budget Neutral; *A+B* = Combined A & B Amount

^cThe minimum required bundled services designated by provider type: *H* = Hospital; *PA* = Physician, all; *PS* = Physician, subset; *PG* = Physician Gainshare; *PAC* = Post-acute care; *O* = Other (includes laboratory, durable medical device suppliers, and/or prescription drugs); *R* = Readmission

^d*Episode duration* is the postdischarge window (+ Preadmission if applicable).

^e*P* = Prospective; *R* = Retrospective; *Q* = Quality-based

^fNumber of hospital participants as of February 20, 2017 and May 17, 2019.

^gCMMI does not provide the number of hospitals that are subject to IPPS, likely because IPPS applies to all nonfederal community acute care hospitals that contract with Medicare. These numbers come from the American Hospital Association's 2019 and 2017 counts of community care hospitals.

^hUnder HBC, the bundle included only hospital and physician services provided during the index event and any readmission within 72 hours.

ⁱCSAP had four participating ambulatory surgical centers (rather than hospitals) and the centers' associated physicians. Ambulatory surgical centers are paid under Medicare Part B. The discount includes both physician professional and ambulatory surgical center facility fees.

^jThe 30-day post-discharge window only applies to related readmissions. All other services provided postdischarge are not included in the episode.

^kOn November 30, 2017, CMS restricted the number to the 34 metropolitan areas that were already participating in the CJR program. Originally, the program was to expand to 67 areas.

^lThe final rulings for AMI, CABG, SHFFT, and CR models were announced December 20, 2016. They were originally scheduled to start in the first half of 2017. However, CMS delayed the start date in March 2017 and eventually canceled the programs on November 30, 2017.

^mThe CR model set prices per services as follows: \$25/service for the first 11 services and \$175/service thereafter.

to CJR and would have required providers in certain metropolitan areas to participate.

Estimates of the impact of a voluntary initiative may be affected by selection bias because the providers that apply and are selected by Medicare may not be representative of all US hospitals. For example, studies have shown that compared to nonparticipants, BPCI Model participants have been more likely to be located in urban areas, be larger both in terms of the number of beds and patient volume, and have shorter lengths of stay.¹⁸⁻²⁰ Nonvoluntary initiatives, such as CJR, may be less susceptible to selection bias due to random assignment of participation. Although mandatory participation with randomization may be preferable for research purposes and generalizability, it may not be achievable for political reasons. In addition, it may put some hospitals at a competitive disadvantage if, for example, they are forced to participate in a payment model for which they are not prepared, while their competitors in the control group operate under the status quo.²¹

Contracting Flexibility

In several pre-ACA initiatives, Medicare solicited applicants to offer terms of participation. Provider applicants offered discounts off IPPS rates to Medicare in their proposals and described how they would achieve improved quality and lower production costs. Participants could expand the scope of the episode by including more types of providers than the minimum set required by Medicare. For example, the minimum set for the HBC initiative included surgeons, anesthesiologists, cardiologists, and radiologists, but some applicants included additional provider types (eg, pulmonologists and neurologists). On the basis of providers' proposed terms, Medicare selected providers to participate.

In some recent initiatives (eg, HG, PHC, and BPCI models), Medicare was less flexible in its contracting parameters. In these programs, unlike in prior initiatives, Medicare defined the services that would be included in the bundle (ie, almost all Part A and B services), and it set a menu of options from which providers could choose, including discount rates up to 3%, and episode durations of 30, 60, or 90 days postdischarge. The choice of duration was sometimes tied to a particular discount rate. For example, in BPCI Model 2, providers choosing a 30- or 60-day duration had to accept a 3% discounted price and a 2% discount for 90-day durations. Applicants could choose from a list of Diagnosis Related Groups

(DRGs) to which the contracts would apply. Providers willing to accept the structure could participate.

As noted previously, participation in the CJR initiative has been mandatory for those providers selected by a random assignment process. In this initiative, Medicare has provided little choice for participants. Episode durations have been set at 90 days, and services have included almost all Part A and B services. Discounts are based on quality metrics. The canceled SHFFT, AMI, CABG, and CR programs were designed similarly to CJR, having less flexible contracting terms than previous models.

Findings from initiatives in which participants have contract flexibility can be limited by selection bias. If providers can craft their bundle (as they did in the HBC initiative), they are more likely to do so in ways that increase their chances of success. This threatens generalizability of the results from these initiatives. Conversely, voluntary and flexible programs are more inviting to providers, potentially facilitating more rapid expansion of bundled payments. Medicare seems to be testing this trade-off: Some recent initiatives have permitted flexibility across preset categories, while others have featured less flexibility and mandatory participation (for those randomly assigned).

By being more prescriptive in initiative design, Medicare can simplify administrative burdens and reduce variety in implementation to generate more consistent empirical results. The specific parameters chosen by Medicare to define the contract may reflect the lessons CMS has learned from past models.

Medicare Payment Rates and Potential Cost Savings

The amount that providers discount their reimbursement from Medicare is likely to affect whether Medicare achieves cost savings. Some initiatives (IPPS, HG, and PHC) were budget neutral, meaning participants had to pay Medicare back if reimbursement exceeded previous reimbursement. In other initiatives, hospitals offered discounts to Medicare; for example, HBC hospital participants offered discounts between 9.7% and 36.7%, the largest discounts of all the initiatives. In recent initiatives (BPCI Models 2-4, BPCI Advanced, and CJR), Medicare has provided a menu of discount rates between 2% and 3%; these minimal discounts may be intended to encourage participation.

Medical Conditions

The medical conditions covered by the Medicare initiatives have varied. Seven initiatives (HBC, ACE, CJR, and the canceled SHFFT, AMI, CABG, and CR models) have focused specifically on episodes initiated by orthopedic and cardiovascular inpatient surgeries. Eight initiatives (IPPS, HG, PHC, BPCI Models 1-4, and BPCI Advanced) have offered more expansive lists of medical conditions from which providers could choose. For example, BPCI model participants could select from 48 diagnoses. Many BPCI Model participants have chosen to focus on orthopedic and cardiovascular procedures. The remaining demonstration (CSAP) focused on cataract surgery.

Medicare has emphasized the inclusion of orthopedic and cardiovascular inpatient surgical episodes in bundled payments for several reasons. These episodes comprise a large fraction of Medicare payments, are high volume, and are believed to have high profit margins for providers. Hip and knee replacement surgeries are the most common inpatient surgeries paid for by Medicare. In 2014, Medicare paid for more than 400,000 of these procedures, with a cost to Medicare of more than \$7 billion for hospitalizations alone.²² Medicare has also focused on bundled payments for the care of orthopedic and cardiovascular surgical patients because these patients typically have a long recovery period that includes rehabilitation and care after discharge from acute care. There is large variation in Medicare spending on post-acute care services for these conditions, and metrics are available to evaluate the effect of the bundled payment demonstrations on quality.

As Medicare has gained experience on bundled payments for orthopedic and cardiovascular surgeries and fine-tuned the design features, it has experimented with bundled payments for a wider range of conditions, such as infectious diseases, gastrointestinal conditions, stroke, and renal failure. However, the expansion has brought to the forefront one concern about bundled payment contracts as they have been implemented thus far: A fixed payment amount per condition might not accurately depict the potential costs of a patient's episode of care. Although several initiatives have set outlier thresholds, these thresholds may not be enough to cover costs. Providers may be incentivized to upcode if they anticipate that the cost of a patient's care will be higher than the bundled payment. The way in which condition groups are risk adjusted is important work for future implementations of bundled payments.

Episode Duration

Longer episode durations can put greater financial risk on providers, since they will be more accountable for patient outcomes and the quality of care further into the future from the episode event. Several demonstrations—IPPS, ACE, and BPCI Model 1—include only the index hospitalization in their episode definition. Other demonstrations, like CSAP, include services prior to cataract surgery through 120 days post-surgery. Even so, the evaluation report for CSAP mentioned a potential issue of providers waiting to provide services until after the 120 days.²³ HG and PHC include services two weeks prior to and 30 days after the hospitalization. BPCI Model 2 and BPCI Advanced include the hospitalization and continue 30, 60, or 90 days post discharge, depending on the participants' choice. BPCI Model 3 is similar, but it excludes the initial hospitalization. The most prescribed and nonvoluntary initiatives—CJR and the four canceled ones—include the hospitalization and services provided within 90 days after discharge.

Services in the Bundle

The types of services included in a bundle definition affect the types of providers that are held accountable for a patient's episode of care. The least comprehensive initiatives (IPPS and BPCI Model 1) have included only inpatient hospital services. More inclusive initiatives (eg, HBC, CSAP, ACE, and BPCI Model 4) have included hospital and physician services. The most inclusive forms of initiatives thus far (BPCI Model 2, CJR, and the canceled SHFFT, AMI, CABG, and CR initiatives) have included nearly all Medicare Part A and B services, except for hospice, readmissions for major trauma, new technology add-on payments, and a few other carve-out services.

Among the Medicare bundled payment initiatives, BPCI Model 3 was unique in both its definition of the episode of care and the services in the bundle. Although the index event in this initiative was a hospital admission, the episode for bundled payment purposes would not begin until care was provided by a post-acute care provider; this care had to begin within 30 days of the related hospital discharge. Post-acute care providers in the initiative included skilled nursing facilities (SNFs), inpatient rehabilitation facilities (IRFs), long-term care hospitals (LTCHs), and home health agencies (HHAs). Similar to BPCI

Model 2, all Part A and B services within the episode were included in the bundle, including readmissions. Unlike BPCI Model 2, post-acute care providers could be the episode applicants/awardees in BPCI Model 3; hospitals, physician practices, and physician hospital organizations were eligible to apply to both BPCI Models 2 and 3.

BPCI Model 3 was more inclusive than the status quo payment systems for post-acute care providers. Typically, Medicare pays a case-mix-adjusted, prospectively determined per diem payment to SNFs for their services, a case-mix-adjusted, prospectively determined payment per 60-day episode to HHAs, and a case-mix-adjusted, prospectively determined payment to LTCHs and IRFs for their services.

More inclusive bundle definitions may encourage providers to branch out from their traditional fee-for-service silos and coordinate with other providers in the bundle. This could result in a different composition of services delivered to patients. The bundle definition could also affect the composition of services if providers shift to services and settings of care that are outside the bundle (in order to save on costs). Coulam and Gaumer report several studies finding that there was a significant shift from inpatient services to outpatient services following the implementation of IPPS, which only included inpatient hospital services in the bundle.²⁴ For example, diagnostic tests and preparations for surgery shifted from being performed in inpatient hospital settings to outpatient settings.²⁵ When the bundle is more expansive and includes both inpatient and outpatient services, the incentive for such a shift is diminished.

Retrospective, Prospective, and Value-Based Payments

The ways that payments have been administered has varied from initiative to initiative. In early demonstrations, Medicare paid participants a prospectively negotiated amount per episode. In more recent initiatives, including many of those initiated under the ACA, payments have been retrospective with a target price. This means that shortly after a patient receives a service, Medicare pays the provider according to its status quo contract, and the payment is subsequently reconciled with the target price for the service. For example, physicians would continue to bill based on the fee-for-service rate and hospitals would continue to bill based on Medicare Severity Diagnosis Related Groups (MS-DRGs),

and then CMS would provide quarterly or annual feedback on how the average amount paid per episode compared to the target price. Providers with average payments exceeding the target would be required to pay back Medicare, up to a certain limit. Those with average payments below the target could keep the difference, up to a certain amount. For example, in BPCI models, gains and losses relative to target prices have been capped at 20% of the target price.

The target price has typically been a discount off historical average payments for a given condition. Sometimes, the prices have depended on actual national average changes in Medicare spending (as in the case of BPCI models) or on annual Medicare payment updates (as in the case of CJR). Retroactively determined target prices create uncertainty for providers, which may lead to a delayed impact of bundled payments or attrition in participation.

In CJR, the discount (and thus the target price) is contingent on quality performance, with smaller discounts rewarding higher quality. Although this payment methodology may delay reconciliation and create more uncertainty for providers, CJR is the first initiative to link reimbursement to quality measures (ie, value-based payments). Thus, we might expect CJR to have larger effects on the quality of care than was associated with initiatives that focused only on spending.

Methods

Literature Review

We focused this review on quantitative studies of Medicare's bundled payment initiatives centered on hospital-initiated episodes. To identify studies, we used the Publish or Perish software by Anne Harzing (version 6 from November 2, 2018) to extract Google Scholar search results of study titles that met at least one of the following criteria:

- All of the words: bundled payment OR bundled payments OR bundle payment OR bundle payments (ie, bundled payment[s] OR bundle payment[s])
- All of the words: bundling payment(s) OR episode payment(s) OR episode based payment(s) OR prospective payment(s); AND any of the words: impact(s), effect(s), association(s), result(s),

Figure 2. Citation Inclusion and Exclusion Categories and Counts

N=1,473 Google Scholar (GS) Search Results	
Pub year <2000	Pub year ≥ 2000
578 citations	895 citations
↓ Exclusion (529)	↓ Exclusion (853)
<=30 cites 529	Duplicate 13
49 citations	Not U.S. 78
↓ Exclusion (28)	Not journal publication 53
Duplicate 2	Other PPS 123
Not U.S. 1	Regulatory 78
Not journal publication 1	Not relevant analysis 134
Other PPS 4	Perspective/editorial 195
Not relevant analysis 19	No abstract/article online 177
Perspective/editorial 1	Textbook/book 2
21 citations	42 citations
↓ Exclusion (11)	↓ Exclusion (25)
IPPS Pre-post 9	Post Data Only 3
IPPS Descriptive/Trends 2	IPPS Pre-Post 3
10 citations	BPCI Pre-Post 15
IPPS DiD 5	BPCI Descriptive/Trends 1
IPPS Predict vs Actual 1	CJR Pre-Post 1
IPPS Review 3	CJR Simulation 1
HBC Various Methods 1 ^a	17 citations
	ACE DiD 1
	BPCI DiD 6 ^b
	CJR DiD/LATE 1
	HBC Pre-Post 1 ^c
	BPCI Pre-Post 1 ^c
	Review 8
N=8 added CMS evaluation reports	N=2 added citations not in GS search
CSAP Descriptive/Qualitative 1	IPPS DiD 1
HG DiD 1	IPPS Review 1
PHC DiD 1	
HBC Various Methods 1 ^a	
ACE DiD 1	
BPCI Model 1 DiD 1	
BPCI Models 2,3,4 DiD 1 ^b	
CJR DiD 1	

Abbreviations: *DiD*, research design is difference-in-differences; *LATE*, local average treatment effect; *Pre-Post*, compares outcomes pre- and post-implementation of initiative; *CMS*, Center of Medicare and Medicaid Services. See appendix for complete list of abbreviations. ^a There were two evaluations of HBC that were published prior to the year 2000: (1) Cromwell et al. (1997),²⁶ a publication in *Health Care Financing Review* (CMS's dissemination platform); and (2) Cromwell and the Health Economics Research Group (1998),²⁷ the official evaluation report to CMS and Congress.^{26,39} The first was a condensed version of the second and examined only the first four participants in HBC,²⁷ and the second examined all seven participants.³⁹ ^b One of the BPCI DiD Google Scholar citations is the 2-year evaluation report of BPCI Models 2-4 by the Lewin Group. We did not use this report; instead, we used their 3-year and 5-year evaluation reports. ^c Although these two studies used a pre-post design, they are included in our review because they provide quantitative evidence on hospital production costs, which is extremely rare in the literature.

evidence, compare(s), comparison(s), comparing, evaluation, evaluating, implication(s), lessons, assess(ing/ment), quality, saving(s), value, outcome(s), or cost(s)

- Comprehensive Care for Joint Replacement

The search resulted in 1,473 citations. We filtered the results for (a) duplicate citations; (b) publications pertaining to non-US bundled payment initiatives; (c) articles published in non-US journals; (d) doctoral dissertations, master's theses, posters, presentations, and unpublished working papers; (e) articles on bundled payment initiatives outside the scope of this study (eg, non-Medicare initiatives and initiatives unrelated to a hospital admission, such as models exclusively for SNF, HHA, or end-stage renal disease services); (f) regulation-related publications (eg, comments or rulings); (g) publications examining outcomes other than cost, quality, or utilization (eg, technology adoption or competition); (h) studies that used weaker empirical methods and/or did not directly estimate the effect of bundled payments (eg, cost analyses predicting hypothetical cost savings for a bundled payment initiative, analyses on the variation in a related outcome variable such as length of stay or hospital costs, analyses on the predictors of cost or utilization, or trend analyses under an initiative without a control group); (i) publications that did not provide any empirical evidence (eg, perspective or guidance articles); and (j) studies that were either unpublished or unavailable online (including two books). See Figure 2 for the results of the filtering process.

To simplify our categorization, we divided studies based on publication date: one group included studies published before 2000; the other included studies published in 2000 or later. We found that studies in the former group were almost exclusively about IPPS (except for one study about HBC); studies in the latter group were about initiatives after IPPS, since most of them were implemented after the year 2000. We found that the literature has focused on a subset of initiatives: HBC, ACE, BPCI, and CJR.

After we applied the inclusion/exclusion criteria, 27 citations remained. We augmented the list with relevant articles cited in the 27 publications, as well as CMS evaluation reports of each initiative produced by contracted research organizations (ie, Abt Associates, Health Economics Research, RTI International, IMPAQ, Econometrica, and the Lewin Group). The evaluation reports, most of which can be found via the CMMI website, tended to be comprehensive and provided both

qualitative and quantitative findings. (Note: Authors of two of the CMS reports have published subsets of their findings in academic journals: Health Economics Research published some of their findings on HBC in *Health Care Financing Review*,²⁶ and the Lewin Group published some of their 1-year evaluation findings on BPCI Model 2 in the *Journal of the American Medical Association*.¹⁸ To distinguish journal articles from evaluation reports, we refer to evaluation reports by the research organization's name, such as Health Economics Research²⁷ or Lewin Group,¹⁹ rather than the name of the first author of the publication.)

In total, we examined 37 studies: 25 original research studies and 12 literature reviews. The studies about IPPS included 4 literature reviews^{12,24,28,29} and 7 original research studies.³⁰⁻³⁵ The research studies involved difference-in-differences analyses or compared predicted values from a "time-series" analysis of pre-IPPS data with actual values under IPPS, thus providing a counterfactual. The studies on other initiatives included 8 reviews^{1-5,36,37} and 18 original research studies.^{13,14,16,18-20,23,26,27,38-46} Most of the original research had higher-quality designs, but one study did not: Although Navathe and colleagues¹³ only performed pre-post analyses without a comparison group, we included their study because it was the only one to provide quantitative evidence on hospital internal production costs under a bundled payment structure that included hospital, physician, and post-acute care services. Table 2 lists each included study, describes the outcomes studied, and grades the study designs for each outcome. (Note: Our tables and figures mostly present results from the original research articles. However, we frequently refer to the review articles in our discussion.)

In our analysis, we focused on estimates produced by higher-quality research designs (rating of A or B+ in Table 2) and excluded studies lacking a control group (eg, pre-post comparison studies). Without a control group or counterfactual, the estimated effect of bundled payments may be biased. Not every outcome variable for every initiative was studied using a research design rated A or B+; therefore, some of the figures present results from fewer studies. Payer spending, length of stay, readmission rates, and post-acute care utilization and spending were the most frequently studied variables across initiatives. Other outcome variables—case-mix selection, volume, and quality measures—were less commonly studied with research designs rated A or B+.

Table 2. Study Design Quality by Outcome Variable in Empirical Studies on Medicare Bundled Payment Initiatives

Article ^a	CMS Report ^b	Eval. Year ^c	Costs		Resources			Postdischarge			Quality		Other	
			Payer	Hospital	LOS	Readm.	ED	PAC	Mortality	Functioning	Complications	Episode Volume	Case Mix	
Hospital-Only Initiatives														
<i>IPPS</i>														
Feder et al. (1987) ³⁰	N		B+	B+	B+									B+
DesHarnais et al. (1987) ³¹	N	1	B+	B+	B+	B+	B+	B+	B+	B+	B+	B+	B+	B+
Guterman et al. (1986) ³²	Y	1	C	Q	B+	C	C	C	C	C	C	C	C	C
Sloan et al. (1988) ³³	N	2	A	A	B+	B+	B+	B+	B+	B+	B+	B+	B+	B+
Mayer-Oakes et al. (1988) ³⁴	N	2	B+	B+	B+	B+	B+	B+	B+	B+	B+	B+	B+	B+
Frank et al. (1986) ³⁵	Y	2	A	A	A	A	A	A	A	A	A	A	A	A
<i>HG</i>														
RTI (2014) ³⁸	Y	3	A	Q	A	A	A	A	A	A	A	A	A	B+
<i>PHC</i>														
RTI (2014) ¹⁶	Y	3	A	Q	A	A	A	A	A	A	A	A	A	B+
<i>BPCI Model 1</i>														
Econometrica (2016) ²⁰	Y	2	A	Q	A	A	A	A	A	A	A	A	A	B+
Hospital+Physician Initiatives														
<i>HBC</i>														
Cromwell et al. (1997) ²⁶	Y	3	B	C	C	D	D	D	D	D	D	D	D	C
Health Economics Research (1998) ²⁷	Y	5	B	C	C,D	B	B	B	B	B	B	B	B	C
Liu, Subramanian, Cromwell (2001) ³⁹	N	2.5	C	C	C	C	C	C	C	C	C	C	C	C

Continued

Table 2. Continued

Article ^a	CMS Report ^b	Eval. Year ^c	Costs		Resources		Postdischarge			Quality		Other	
			Payer	Hospital	LOS	Readm.	ED	PAC	Mortality	Functional- omg	Comple- ations	Episode- Volume	Case Mix
ACE													
IMPAQ (2013) ¹⁴	Y	3	B	A	A	A	A	A	A	A	C	A	A
Chen et al. (2018) ⁴¹	N	2	A		A	A	A	A	A	A	A		
BPCI Model 4													
Lewin Group (2018c) ⁴⁰	Y	3		Q	A	A	A	A	A	A	A	C	
Hospital+Physician+Post-Acute Care Initiatives													
CSAP													
Abt (1997) ²³	Y	3	C						C		C	C	C
BPCI Model 2													
Lewin Group (2018b) ¹⁹	Y	5		Q	A	A	A	A	A	A	D		C
Lewin Group (2018c) ^{40,4}	Y	3									A		
Dummit et al. (2016) ¹⁸	N	1.66		Q	A	A	A	A	A	A	A	A	A
Joynt Maddox et al. (2018) ⁴²	N	2	A		A	A	A	A	A	A		A	A
Martin et al. (2018) ⁴³	N	1	A						A			A	A
Navathe et al. (2018) ⁴⁹	N	2.25										A	A
Jubelt et al. (2017)	N	1	A		A	A	A	A	A	A			
Navathe (2017) ¹³	N	2	C	C	C	C	C	C	C	C		C	C
CJR													
Finkelstein et al. (2018) ⁴⁵	N	1	A		A	A	A	A	A	A		A	A
Lewin Group (2018a) ⁴⁶	Y	1	A		A	A	A	A	A	A			

Continued

Table 2. Continued

Article ^a	CMS Report ^b	Eval. Year ^c	Costs		Resources		Postdischarge			Quality		Other		
			Payer	Hospital	LOS	Readm.	ED	PAC	Mortality	Functioning	Complications	Episode Volume	Case Mix	
Post-Acute Care Initiatives Only (within 30 days of hospitalization)														
BPCI Model 3														
Lewin Group (2018b) ¹⁹	Y	5		Q	A	A	A	A	A	D		C		
Lewin Group (2018c) ^{40,41}	Y	3								A		C		
IPPS Reviews														
Coulam and Gaumer (1992) ²⁴	N		R	R	R	R	R	R	R	R		R	R	
Rosenberg and Browne (2001) ²⁹	N		R	R	R	R	R	R	R				R	
Lave (1989 & 1990) ^{9,28}	N		R	R	R	R	R	R	R	R		R	R	
Feinglass and Holloway (1991) ¹²	N		R	R	R	R	R	R	R	R		R	R	
Bundled Payment Reviews														
Beyond IPPS														
Hung et al. (2018; systematic review) ⁴	N		R					R						
Siddiqi et al. (2017; systematic review) ³	N		R		R		R							
Hussey et al. (2012; systematic review) ⁵	N		R	R	R	R	R	R	R	R		R	R	
Glickman, Dinh, Navathe (2018) ^{36,41}	N		R		R	R	R	R	R	R		R	R	

Continued

Table 2. Continued

Article ^a	CMS Report ^b	Eval. Year ^c	Costs		Resources			Postdischarge			Quality		Other	
			Payer	Hospital	LOS	Readm.	ED	PAC	Mortality	Functional- outcomes	Complic- ations	Episode Volume	Case Mix	
McLachhorn and Buller (2017) ¹	N		R		R	R	R	R			R			
Piccinin et al. (2017) ³⁷	N		R		R	R	R							
Shih et al. (2015) ²	N		R		R	R	R	R			R			

See appendix or Table 1 for list of initiative abbreviations. *Study design quality*: Grade A analyses have a control group and use longitudinal variation, such as difference-in-differences (DiD) analysis. Grade B+ analyses provide enough data to perform a DiD analysis but do not compute a DiD estimate; typically, the analysis is descriptive, meaning no control variables were included and standard errors were not provided. Grade B analyses compare a computed counterfactual of an outcome variable based on trends and the actual measure of the variable. Grade C analyses do not have a separate control group and compare episodes in the pre-period with episodes during the bundled payment initiative. Grade D analyses show cross-sectional differences in outcome variables between episodes from participating hospitals with that of comparison hospitals during the initiative period. Q represents qualitative findings, such as from interviews or focus groups. R represents a literature review.

^aY indicates that the study either is published in *Health Care Financing Review* (CMS's dissemination platform from 1979 to 2009) or is commissioned by CMS and performed by a consultancy group.

^b*Evaluation year* indicates the number of years between the initiative start date and the last episode studied.

^cThe Lewin Group 3- and 5-year evaluation reports of BPCI Models 2–4 are similarly structured. We have reported findings from whichever report provides the best available estimate for a given measure, prioritizing the research design quality and the higher evaluation year. For some measures (patient functioning) and for BPCI Model 4, the 3-year report provides DiD estimates, but the 5-year report does not.

^dGlickman, Dinh, Navathe (2018)³⁶ is an issue brief and was not peer reviewed.

Most studies in our review examined the effectiveness of bundled payment initiatives on specific medical conditions. Typically, they focused on between one and five conditions. However, the Lewin Group 5-year evaluation of BPCI Models 2-4,¹⁹ included results for 34 medical conditions. For brevity, we focused on the 14 conditions in that review that overlapped with the medical conditions studied by other investigators so that we had a sample size of two or more studies for each condition: major joint replacement of the lower extremities (MJRLE), major joint replacement of the upper extremities (MJRUE), hip or knee revisions, spinal fusions, CABG, AMI, cardiac valve surgery, percutaneous coronary intervention, congestive heart failure (CHF), defibrillator implant surgery, chronic obstructive pulmonary disease (COPD) and other related issues, urinary tract infection, pneumonia, and sepsis.

We summarized the impact of bundled payment models on the following 12 measures: Medicare payments to providers, hospital internal production costs to providing services, length of stay, the likelihood of using post-acute care, Medicare payments to post-acute care providers, readmission rates, mortality rates, emergency department (ED) use, adverse patient outcomes, patient functioning, episode volume, and case-mix selection. The figures show the effects in absolute terms. See the Online Supplemental Material for figures showing the effects in percentage terms.

For most measures, we focused on quantitative evidence from higher-quality research designs. However, we made one exception to this approach: For hospital production costs, we additionally discussed qualitative findings because the quantitative evidence was limited.

Table and Figure Design

The figures present most of the findings on bundled payments (except for hospital internal production costs). For a given measure (eg, Medicare spending per episode), we plotted the point estimates and 95% confidence intervals (CIs) or 1.96-times the standard error (SE) whenever possible. In two instances (the BPCI Model 2 study by Joynt Maddox and colleagues⁴² and the Lewin Group report on CJR⁴⁶), the studies did not provide CIs or SEs but there were sufficient data (coefficient estimates, *P* values, and sample sizes) to derive the 95% CIs using the formula $ABS[\text{coefficient}/T.INV.2T(P \text{ value}, N)]$. Cromwell

and associates²⁶ and the Health Economics Research report²⁷ provided estimates for each participant but did not include SEs or measures of statistical significance. For these publications, we computed the average of participants' estimates, and the error bars in the figures represent the highest and lowest estimates among the participants. IMPAQ¹⁴ provided odds ratios for certain measures (eg, mortality rates); we computed the percentage point change so that the estimates were comparable to other studies' estimates. For the relevant measures and corresponding figures, this is noted in the caption of the figure.

We used shorthand labeling for each estimate from the literature. In figures showing estimates for only one measure, such as Medicare spending per episode, the format of the labels is *bundled payment initiative-medical condition studied-first author {or organization name for CMS-contracted reports}-(year of publication)* plus the reference number. In figures that show estimates for more than one measure, such as the measures for case mix, the format of the labels is *bundled payment initiative-medical condition studied-measure studied-first author {or organization name for CMS-contracted reports} (year of publication)* plus the reference number.

We divided the initiatives into four categories based on inclusiveness of the bundle definition: hospital services only (HOS only); hospital plus physician services (HOS+Phys); hospital, physician, and post-acute care services (HOS+Phys+PAC); and post-acute care services only (PAC only). Within each category, we sorted estimates from largest to smallest.

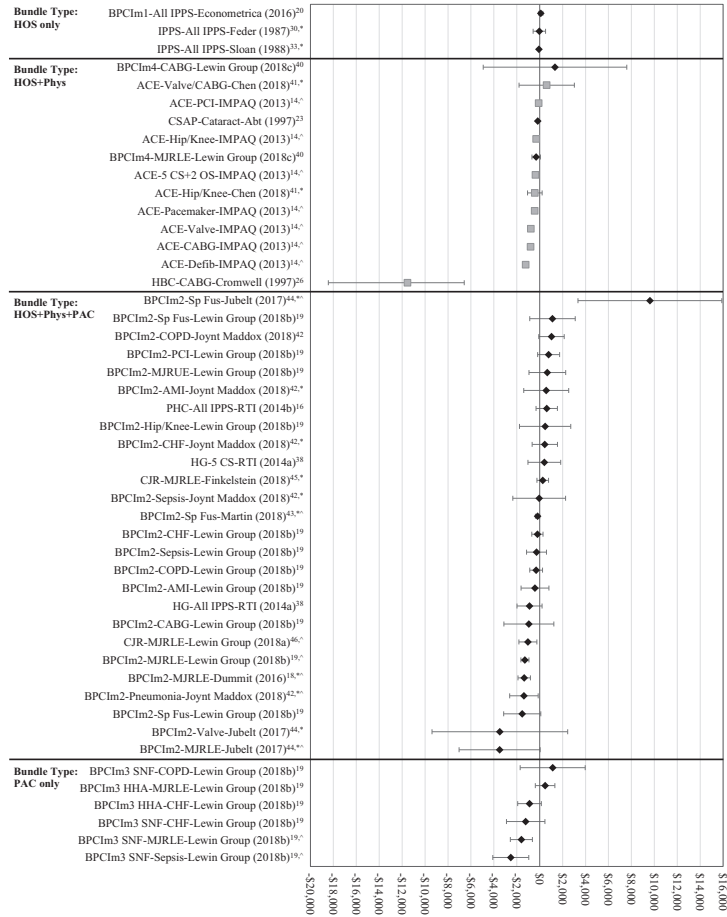
For BPCI Model 3, the post-acute care bundled payment initiative, we typically separated the estimated effects of the initiative by the provider type that was the bundled payment awardee, and we adjusted the first label component to reflect the awardee type. For example, we used BPCIm3 SNF and BPCIm3 HHA for labels of estimates of the model's effects on SNFs and HHAs, respectively.

Results

Payer Spending

Compared with fee-for-service payment models, bundled payment models are expected to contain payer spending. Figure 3 shows inflation-adjusted point estimates of the effect of bundled payment initiatives on

Figure 3. Bundled Payment Impact on Medicare Spending per Episode



Gray squares signify bundled payment initiatives for which a >5% discount off the episode payment was negotiated between Medicare and at least one participant. Black diamonds signify a ≤ 5% discount. Study estimates and 95% confidence intervals are shown with the exception of HBC-CABG-Cromwell (1997),²⁶ which provides estimates for each participant but no standard errors. The figure marks the average of the participants' estimates, and the error bars represent the highest and lowest participant estimate. * indicates peer-reviewed journal publication. ^ indicates estimate was significant at the 5% level. See appendix for list of initiative and condition abbreviations.

Medicare payments to providers with the 95% CIs. Cost estimates were adjusted for inflation using the US Bureau of Labor Statistics Medical Consumer Price Index to be the equivalent of prices in the first half of 2019. The figure shows that 32 of the 48 estimates were negative; however, only a few were significant.

Initiatives in which providers offered or agreed to large discounts (typically the pre-ACA initiatives) yielded larger reductions in payer spending. For example, hospitals in the HBC demonstration offered 9.7% to 36.7% discounts off the combined IPPS Part A amount and physician Part B amount to Medicare, and Medicare's cost savings were between \$3,000 and \$8,500 per episode.²⁶ Hospitals that were part of the ACE demonstration offered discounts for cardiovascular procedures of 0.9% to 8.25% and 0.0% to 4.4% for Part A and Part B services, respectively; and discounts for orthopedic procedures in the ranges of 2.5% to 4.4% and 0.0% to 4.4% for Part A and B services, respectively.^{14,15} Thus, Medicare saved \$585 per inpatient hospitalization in the ACE demonstration,^{15,41} which is substantially less than it saved in the HBC demonstration. In the more recent initiatives (BPCI and CJR), the discount is up to 3% for Part A and B combined, which is smaller than many of the earlier discounts. However, if providers reduce costs to Medicare by more than the agreed upon discount (eg, 3%), they get a share of the savings to Medicare.

A few estimates were statistically significant. Spending (without accounting for reconciliation payments) on multiple joint replacement of the lower extremity (MJRLE) episodes was reduced under several bundled payment models, as evidenced by the Lewin Group's 1-year and 5-year evaluations of BPCI Model 2 and Model 3^{18,19} and the 1-year evaluation of CJR by Finkelstein and colleagues.⁴⁵ However, Finkelstein and colleagues⁴⁵ estimated (using an instrumental variables approach) that CJR led to an increase in Medicare spending when reconciliation payments were included; this finding was not statistically significant. In a study of BPCI Model 2, Jubelt and coauthors⁴⁴ found that participation was associated with a statistically significant increase in spending for spinal fusion episodes. The Lewin Group¹⁹ found significant reductions in Medicare spending for sepsis episodes under BPCI Model 3 when a SNF was the awardee. The study determined that large reductions in post-acute care spending were the main driver for these overall spending reductions.

Hospital Costs and Lengths of Stay

Although not as often studied, the impact of bundled payments on the internal costs of hospitals is important. If bundled payments reduce hospital internal production costs, payers may be able to reduce future health care expenditures. Also, when providers are able to lower production costs, that encourages other providers to participate. This is especially relevant in light of the attrition issues experienced by recent initiatives, such as the BPCI models (see Table 2). Reducing production costs was a priority for many participating hospitals, especially in the early initiatives, and many organizations achieved their goal through gainsharing. When initiatives include post-acute care services, reducing production costs in post-acute care settings may be a priority for participants.

Table 3 provides qualitative and quantitative findings on production costs. Many findings were based on interviews with participating providers, and only some used provider cost data. Note that studies of production costs, unlike the other measures, typically did not have a control group or counterfactual, with the exception of the studies by Feder³⁰ and Sloan,³³ who investigated IPPS and used a difference-in-differences approach.

Table 3 shows that in many initiatives, hospitals reduced internal production costs or slowed the growth rate for those costs. We reported the findings in a table rather than figure because the literature on production costs was relatively heterogeneous in both the methods used (ie, qualitative, pre-post, and difference-in-differences designs) and the ways that production costs were measured (eg, by department or totaled over several years).

Feder³⁰ found that production costs grew more slowly at hospitals paid through IPPS (7.6%) than at hospitals not paid through IPPS (18.1%). Other studies found that production costs also decreased for participants in HG, PHC, HBC, ACE, CSAP, and BPCI Model 2. Three of the four hospitals participating in the HBC demonstration reported reductions between 6.7% and 23.4%; the fourth reported an increase of nearly 11%.²⁶ In the CSAP demonstration, one of the four ambulatory surgical center participants reported reductions in internal costs, another reported increases, and the other two reported no change.⁴⁰ Navathe and colleagues¹³ studied one Texas health care system that participated in

Table 3. Bundled Payment Initiatives and Hospital Internal Production Costs

Study Name ^a	Analysis	Analysis Design and Measures Studied	Results
<i>Hospital-only Initiatives</i>			
IPPS			
IPPS-All-Feder (1987) ³⁰	DiD ^b	Compared pre-post (1982–1984) % changes in production cost per case and in costs among hospitals paid under IPPS and that of hospitals that were given waivers and paid under TEFRA.	Cost per case, IPPS hospitals: +7.6% (Base = \$3,038); Cost per case, non-IPPS hospitals: +18.1% (Base = \$3,079); Total cost, IPPS hospitals: +12.9% (Base = \$17.8 million); Total cost, non-IPPS hospitals: +15.6% (Base = \$21.2 million)
IPPS-All-Sloan (1988) ³³	DiD	Compared pre-post changes in labor, nonlabor, and total production costs per admission among Medicare patients in nonwaiver IPPS states and Medicare patients in waiver IPPS states.	Total hospital cost per admission: +0.4% (not s.s.); Labor cost per admission: –1.7%; Nonlabor cost per admission: +5.0%
BPCIm1-All-Econometrica (2016) ²⁰	Qualitative	Summarized results from interviews with hospital managers.	Significant hospital attrition issues, due to lack of physician engagement and organizations not being able to lower internal costs enough
<i>Hospital + Physician Initiatives</i>			
HBC			
HBC-CABG-Cromwell (1997) ²⁶	Pre-Post	Compared 4 hospital participants' total cost per discharge in 1990 (year before HBC) and that in 1993 (last year of HBC). No information on statistical significance was provided.	SJM Ann Arbor Participant: DRG 106: –\$6,435 (–23.0%); SJ Atlanta Participant: DRG 106: –\$1,910 (–8.6%); BU Boston Participant: DRG 106: –\$2,225 (–6.7%); OSU Columbus Participant: DRG 106: +\$2,773 (+10.9%); SJM Ann Arbor Participant: DRG 107: –\$370 (–2.0%); SJ Atlanta Participant: DRG 107: –\$2,296 (–12.9%); BU Boston Participant: DRG 107: –\$850 (–4.0%); OSU Columbus Participant: DRG 107: +\$4,978 (+24.2%)

Continued

Table 3. *Continued*

Study Name ^a	Analysis	Analysis Design and Measures Studied	Results
HBC-CABG-Health Economics Research (1998) ²⁷	Pre-Post	Compared 3 hospital participants' direct cost per discharge for a given cost center in 1990 (pre-HBC) and that in 1993 (post-HBC). Overhead in the direct costs were excluded. Only results significant at the 5% level are reported here. The number of hospitals for which results were statistically different from 0 are listed in parentheses.	Nursing ICU: DRG 106: -34% to -25% (2); Nursing general: DRG 106: -40% to -32% (2); Pharmacy: DRG 106: -35% to -32% (2); Operating room: DRG 106: -16% to +19% (2); Catheter lab: DRG 106: -29% (1); Respiratory therapy: DRG 106: (0); Radiology: DRG 106: -25% (1); Laboratory/blood bank: DRG 106: -45% to +39% (2)
HBC-CABG-Liu (2001) ³⁹	Post-only	Compared 3 hospital participants' direct cost per discharge in 1991 (first year of HBC) and that in 1993 (last year of HBC). Controlled for case mix. Overhead in the direct costs were excluded. All results were statistically significant at the 5% level.	SJM Ann Arbor Participant: DRG 106: -\$3,942 (-30%); SJ Atlanta Participant: DRG 106: -\$2,517 (-19%); BU Boston Participant: DRG 106: -\$1,824 (-14%); SJM Ann Arbor Participant: DRG 107: -\$2,678 (-27%); SJ Atlanta Participant: DRG 107: -\$1,848 (-18%); BU Boston Participant: DRG 107: -\$1,124 (-12%)
	Post-only	Compared 3 hospital participant's direct cost per discharge for a given cost center in 1991 (first year of HBC) and that in 1993 (last year of HBC). Controlled for case mix. Overhead in the direct costs were excluded. Only results statistically significant at the 5% level are reported here. The number of hospitals for which results were statistically different from 0 are listed in parentheses.	Routine Nursing: DRG 106: -55% to -21% (3); Nursing ICU: DRG 106: -41% to -31% (3); OR and recovery: DRG 106: -33% to +6% (2); Pharmacy: DRG 106: -38% to -32% (2); Catheter lab: DRG 106: -52% to -31% (2); Other: DRG 106: -33% to -15% (2); Nursing ICU: DRG 107: -56% to -22% (2); OR and recovery: DRG 107: -38% to -29% (3); Pharmacy: DRG 107: -32% to +6% (2); Other: DRG 107: -36% (1)

Continued

Table 3. *Continued*

Study Name ^a	Analysis	Analysis Design and Measures Studied	Results
ACE-CS-IMPAQ (2013) ¹⁴	Qualitative	Provided changes in total hospital admission internal costs for 4 hospital participants (as self-reported by hospital managers). We divided this number by the number of episodes to capture the per episode change in internal costs.	HMC participant: -\$481; BHS participant: -\$1,355; ESJH participant: -\$5,606; OHH participant: \$0
ACE-OS-IMPAQ (2013) ¹⁴	Qualitative	Provided changes in total hospital admission internal costs for 3 hospital participants (as self-reported by hospital managers). We divided this number by the number of episodes to capture the per episode change in internal costs.	HMC Participant: -\$609; BHS Participant: -\$2,558; LHS Participant: -\$696
ACE-CS/OS-IMPAQ (2013) ¹⁴		Provided gainsharing amounts hospitals paid to cardiovascular and/or orthopedic physicians per episode (as self-reported by hospital managers).	BHS Participant: \$350; HMC Participant: \$320-\$390; LHS Participant: \$303; ESJH Participant: \$597
BPCI _m 4-CS/OS-Lewin Group (2018b) ¹⁹	Qualitative	Compared the number of hospital participants at the beginning and end of the 2nd year of BPCI Model 4.	Significant hospital attrition issues
CSAP-Catact-Abt (1997) ²³	Qualitative	Interviewed ASC managers about whether they thought they achieved cost savings.	Participant 1: Decrease in costs; Participant 2: Increase in costs; Participant 3: Neutral/no change in costs; Participant 4: Neutral/no change in costs
Hospital+Physician+Post-Acute Care HG-AII-RTI (2014a) ³⁸	Qualitative	Summarized results from focus groups with hospital participant (BIMC) managers.	BIMC reported \$25.1 million in savings over 3 years, with \$2 million in gainsharing payouts to physicians. BIMC initially wanted to reduce device and pharmacy expenditures as well but ultimately that approach was not as successful as focusing on length of stay.

Continued

Table 3. *Continued*

Study Name ^a	Analysis	Analysis Design and Measures Studied	Results
HG-CS-RTI (2014a) ³⁸	Qualitative	Summarized results from focus groups with hospital participant (CAMC) managers.	Through device price negotiations (directly with suppliers or through physicians who had relationships with suppliers), saved \$832,000 for the years of participation. CAMC initially wanted to reduce utilization as well but ultimately did not pursue that path to cost savings.
PHC-All-RTI (2014b) ¹⁶	Qualitative	Summarized results from focus groups with hospital managers.	Significant hospital attrition, likely due to the organizations not being able to lower internal costs enough. Hospitals that were successful reported decreases in internal costs. Sources of hospital cost savings: shifting of services to outpatient, bulk purchasing of drugs, implementation of more cost-effective drugs/services, fewer duplicative diagnostic tests, negotiation of better device pricing, and more timely discharge planning.
BPCIm2-MJRLE-Navathe (2017) ¹³	Pre-Post	Compared one hospital participant's (BHS) internal costs per episode in 2011 (post-ACE, pre-BPCI participation) and 2015 (end of 2nd year under BPCI)	-\$675.11 (-7%) per episode

Select abbreviations: *CS*, Cardiovascular surgery; *All*, All IPCS-eligible inpatient admissions; *ICU*, intensive care unit; *OR*, operating room; *DRG*, diagnosis related group; *SJM*, St. Joseph's Mercy Hospital; *SJ*, St. Joseph's Hospital; *BU*, Boston University Hospital; *OSU*, Ohio State University Hospital; *HMC*, Hillcrest Medical Center; *BHS*, Baptist Health System; *EsJH*, Exempla Saint Joseph Hospital; *OHH*, Oklahoma Heart Hospital; *LHS*, Lovelace Health System; *CAMC*, Charleston Area Medical Center; *BIMC*, Beth Israel Medical Center, ASC, Ambulatory surgical center. See appendix for complete list of abbreviations.

^a*Study Name* denotes the bundled payment initiative, the conditions studied, the first author, and the year of publication.

³⁸Feder et al. (1987)³⁰ did not explicitly provide the difference-in-differences estimate but provided pre-post differences in the control group of hospitals and pre-post differences in the treatment group. We computed manually the difference-in-differences estimate.

both the ACE and BPCI Model 2 initiatives and found that the system reduced hospital internal costs by 20%.

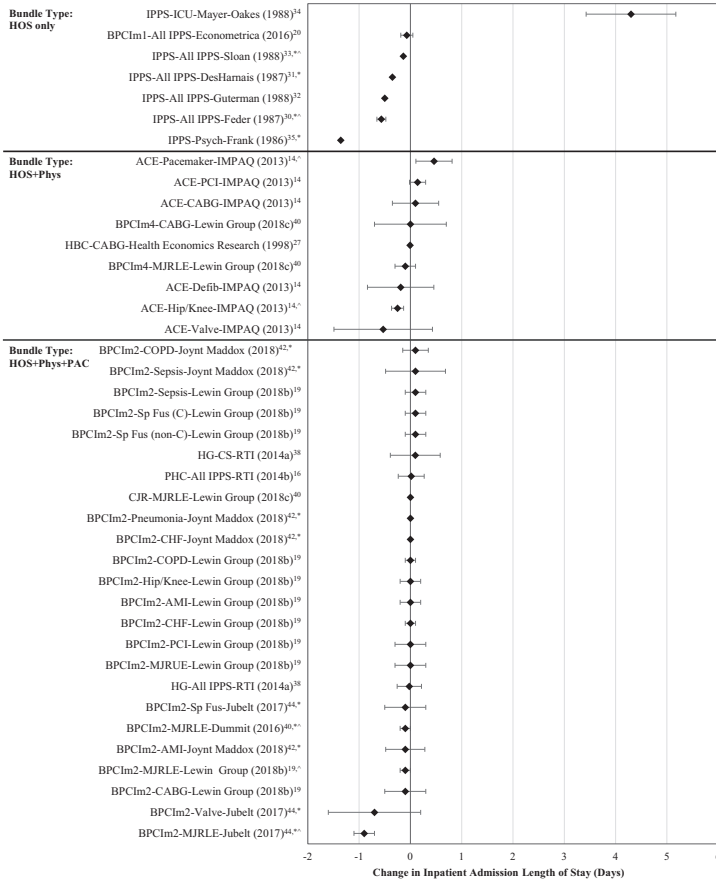
Internal cost reductions came mostly from changes in practices in intensive care units (ICUs), nurse staffing, and the use of laboratory and pharmaceutical products, as well as efforts to negotiate better prices for supplies and durable medical equipment.^{13,14,26} During the ACE demonstration (a HOS+Phys model), some hospitals reported reductions in supply costs and the cost of implants.¹⁴ Although information about the ACA-promulgated initiatives is limited, Navathe and coauthors¹³ reported that under BPCI Model 2 (a HOS+Phys+PAC model), roughly half of the hospital savings came from internal cost reductions (such as implant price reductions), with the remainder coming from reductions in post-acute care costs (eg, by decreasing use of IRFs and SNFs).

Some participants in bundled payment programs reduced internal production costs per episode by shortening length of stay for inpatient admissions (Figure 4). Often, the change in length of stay was not significantly different from that of comparison providers, and in a couple of cases, length of stay either increased in the participating hospitals more than it did in the comparison groups (ACE pacemaker episodes) or the reduction was less in participating hospitals than it was in the comparison groups (IPPS ICU admissions). Under BPCI Model 2, length of stay for MJRLE episodes decreased by 0.1 days¹⁸ and 0.9 days relative to the comparison group.⁴⁴ Under ACE, length of stay for hip and knee replacement episodes decreased by 0.25 days.¹⁴ Due to IPPS, studies^{30-33,35} estimated that the length of stay for inpatient admissions (with the exception of Mayer-Oakes³⁴ estimate for ICU patients) decreased by 0.07 to 1.4 days (an average of 0.6 days across the studies), which is rather large considering this is the average across all inpatient admissions.

Post-Acute Care Service Utilization and Medicare Spending on Post-Acute Care Services

Some bundled payment models (including many but not all recent initiatives) have included post-acute care services in the bundle, and others (typically, the earlier initiatives) have not. One concern among payers and policymakers has been that bundled payment programs could increase the use of post-acute care services as hospitals try to reduce their

Figure 4. Bundled Payment Impact on Length of Stay



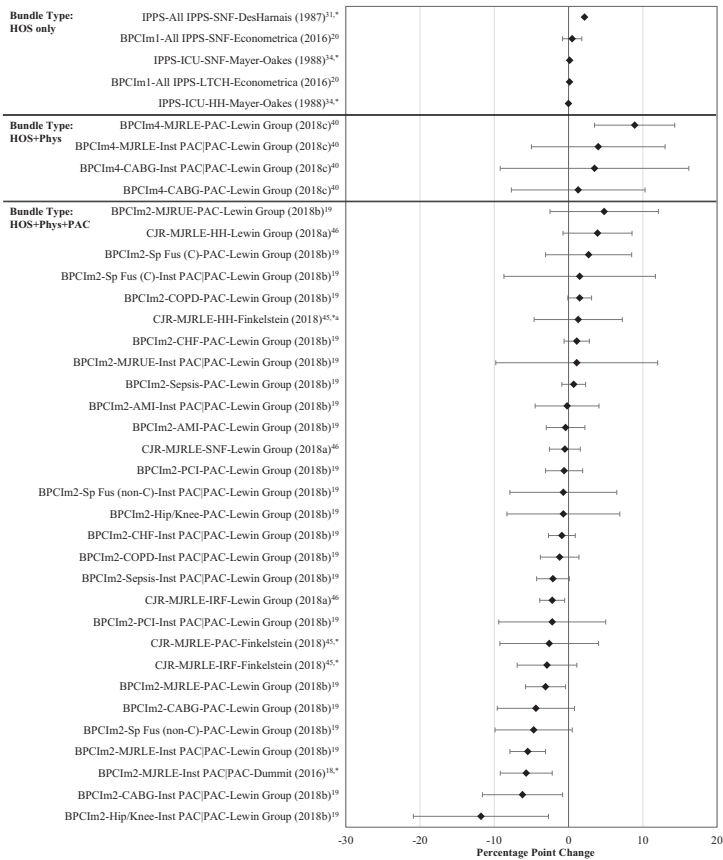
Study estimates and 95% confidence intervals are shown with the exception of HBC-CABG-Health Economics Research (1998),²⁷ which provides estimates for each participant but no standard errors. The figure marks the average of the participants' estimates, and the error bars represent the highest and lowest participant estimate. * indicates the study was published in a peer-reviewed journal. ^ indicates that the results were significant at the 5% level. See appendix for list of initiative and condition abbreviations.

costs and services by discharging patients sooner. Including post-acute care services in the bundled payment model is intended to ensure that hospitals are mindful of the costs of such services and have an incentive to help control them. Hospitals participating in bundled payment models that include post-acute care services have reported that reducing post-acute care use was a priority.¹⁹ It was one of the largest opportunities to reduce spending that did not affect the hospital directly.

Figure 5 shows the estimated effects of bundled payment initiatives on the likelihood of being discharged to or receiving services in various settings. The figure demonstrates that some initiative-condition combinations were associated with increases in the likelihood of post-acute care services, and others were associated with decreases, depending on whether post-acute care services were included in the bundle definition. When post-acute care services were included in the bundle (HOS+Phys+PAC initiatives), 20 of the 29 estimates were negative, whereas eight of nine estimates were positive when the services were excluded (HOS only and HOS+Phys initiatives). However, most of the effects shown in Figure 5 were not statistically different from zero. A few exceptions were the decreases in the use of institutional post-acute care services (given use of post-acute care services) seen for (a) MJRLE episodes under CJR and BPCI Model 2; (b) knee and hip replacements under BPCI Model 2; and (c) CABG episodes under BPCI Model 2. These reductions were rather large, on the order of 9% to 18% (Online Appendix Figure 3). In contrast, the likelihood of using post-acute care services increased for patients treated for MJRLE episodes under BPCI Model 4, a model that did not include post-acute care services in the bundle definition.

With regard to Medicare spending, Figure 6 shows inflation-adjusted estimated effects of bundled payment models on total Medicare payments for post-acute care services (Figure 6a), Medicare payments to IRFs (Figure 6b), Medicare payments to HHAs (Figure 6c), and Medicare payments to SNFs (Figure 6d). We found few estimates on total post-acute care spending, but two episodes (MJRLE and cardiac valve surgery) for a particular hospital studied by Jubelt and associates⁴⁴ showed large cost reductions. In 13 of the 19 estimates on IRF payments, bundles that included post-acute care services in the bundle definition (HOS+Phys+PAC and PAC only initiatives) were associated with cost reductions. However, only three of these associations were

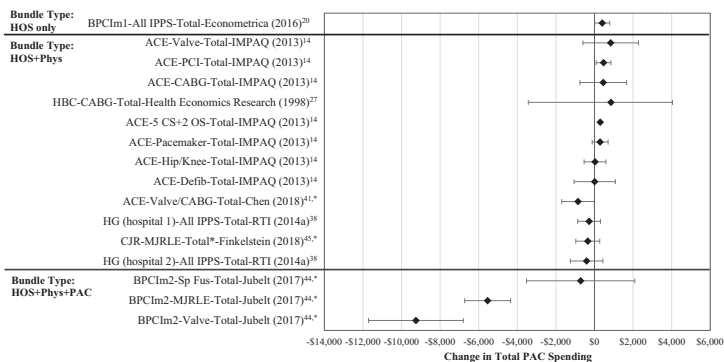
Figure 5. Bundled Payment Impact on the Likelihood of Post-Acute Care Utilization



Study estimates and 95% confidence intervals are shown. * indicates peer-reviewed journal publication. ^ indicates estimate was significant at the 5% level. See appendix for list of initiative and condition abbreviations. The third component of the label denotes the corresponding likelihood measure: *PAC* is the likelihood of patient episodes having any post-acute care services, *HH* of having home health agency services, *SNF* of having services from a skilled nursing facility, *IRF* of having services from an inpatient rehabilitation facility, and *LTCH* of having services from a long-term care hospital. *Inst PAC|PAC* is the likelihood of patient episodes with some PAC services having services from a SNF, IRF, or LTCH.

^a Finkelstein et al. (2018)⁴⁵ estimated the change in the likelihood of ‘no PAC’ for MJRLE episodes under CJR. We negated the sign to be consistent with the other measures.

Figure 6a. Bundled Payment Impact on Total Post-Acute Care Spending

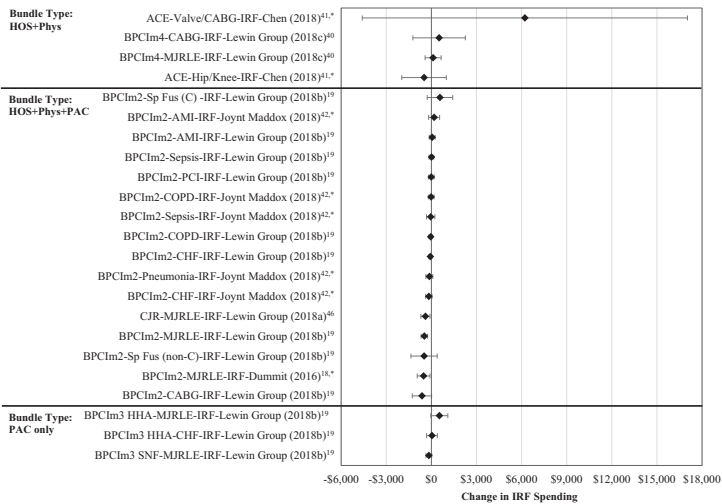


Study estimates and 95% confidence intervals are shown with the exception of HBC-CABG-Health Economics Research (1998),²⁷ which provides estimates for each participant but no standard errors. The figure marks the average of the participants' estimates, and the error bars represent the highest and lowest participant estimate. * indicates peer-reviewed journal publication. ^ indicates estimate was significant at the 5% level. See appendix for list of initiative and condition abbreviations.

statistically significant; they were all for MJRLE episodes under CJR and BPCI Model 2.

Payments to SNFs were often reduced in initiatives that included post-acute care services in the bundle definition. In particular, reductions in the range of \$300 to \$900 (approximately 6% to 13%) were found for CHF, sepsis, MJRLE, and CABG under BPCI Model 2.²⁰ When a SNF was the episode awardee, BPCI Model 3 was associated with approximately \$2,000 per episode (9%-14%) reductions in SNF spending for sepsis, CHF, and MJRLE episodes.²⁰ Nineteen of the 24 estimates were negative for bundles that included post-acute care services in the bundle definition (HOS+Phys+PAC and PAC only initiatives); however, not all findings were statistically significant. In contrast, one of five estimates were negative for bundles that did not include post-acute care services in the bundle definition (HOS only and HOS+Phys initiatives).

Figure 6b. Bundled Payment Impact on Post-Acute Care Spending for Institutional Rehabilitation Facility Services

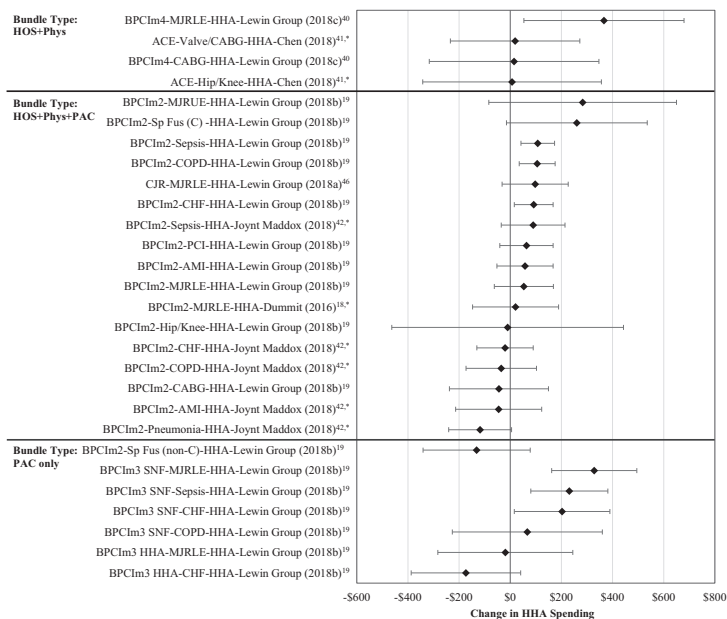


Study estimates and 95% confidence intervals are shown. * indicates peer-reviewed journal publication. ^ indicates estimate was significant at the 5% level. See appendix for list of initiative and condition abbreviations.

Bundled payment initiatives that included post-acute care services (such as BPCI Models 2 and 3) were associated with increased Medicare payments to HHAs for certain conditions: Under BPCI Model 2, the conditions were sepsis, COPD, and CHF; under BPCI Model 3, when a SNF was the episode awardee, the conditions were MJRLE, sepsis, and CHF (Figure 6d). In general, most episodes were associated with increased payments to HHAs (11 of the 18 in the HOS+Phys+PAC group, 4 of the 6 in the PAC only group, and 4 of the 4 in the HOS only group); however, the findings in many studies were not statistically significant.

Contracts that included post-acute care services in episode definitions typically were associated with reductions in the intensity of those services (not shown). For example, BPCI Model 2 was associated with a reduction in the number of days patients stayed at SNFs and an increase

Figure 6c. Bundled Payment Impact on Post-Acute Care Spending for Home Health Agency Services

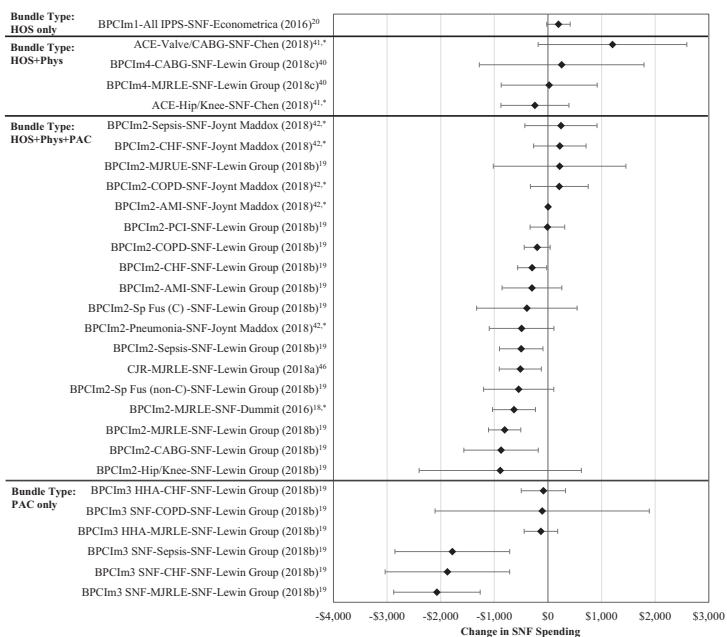


Study estimates and 95% confidence intervals are shown. * indicates peer-reviewed journal publication. ^ indicates estimate was significant at the 5% level. See appendix for list of initiative and condition abbreviations.

in the number of days of home health care.^{13,19} BPCI Model 3 was associated with decreases in the length of stay at SNFs and IRFs.¹⁹ Models that did not include post-acute care services (eg, IPPS) typically were associated with increased use of these services.^{11,12,32,47}

The effect of bundled payments on a particular post-acute care setting seems to depend on the status quo payment system. The status quo for SNF is a per diem rate, for HHA a rate per 60-days of usage, and for IRF a per admission rate. In 2015, Medicare spending per beneficiary was approximately \$21,500 for IRF care, \$17,500 for a SNF stay, and \$5,200 for HHA services.⁴⁸ To reduce Medicare spending on IRFs,

Figure 6d. Bundled Payment Impact on Post-Acute Care Spending for Skilled Nursing Facility Services



Study estimates and 95% confidence intervals are shown. * indicates peer-reviewed journal publication. ^ indicates estimate was significant at the 5% level. See appendix for list of initiative and condition abbreviations.

there would need to be a decrease in IRF admissions. To reduce SNF spending, SNFs would need to reduce the length of patient stays. To reduce HHA spending, there would need to be fewer 60-day episodes. Furthermore, Medicare may spend less on HHA services for the average patient episode than it would for care in a SNF or IRF. With bundled payment incentives, providers may be encouraged to shift patients toward HHA services from the more expensive SNF or IRF services. Figure 5 shows that of all the post-acute care settings, IRF use was the most consistently reduced with bundled payments. Figure 6 shows

that for episodes in which SNF spending was reduced, HHA spending increased.

Readmission Rates

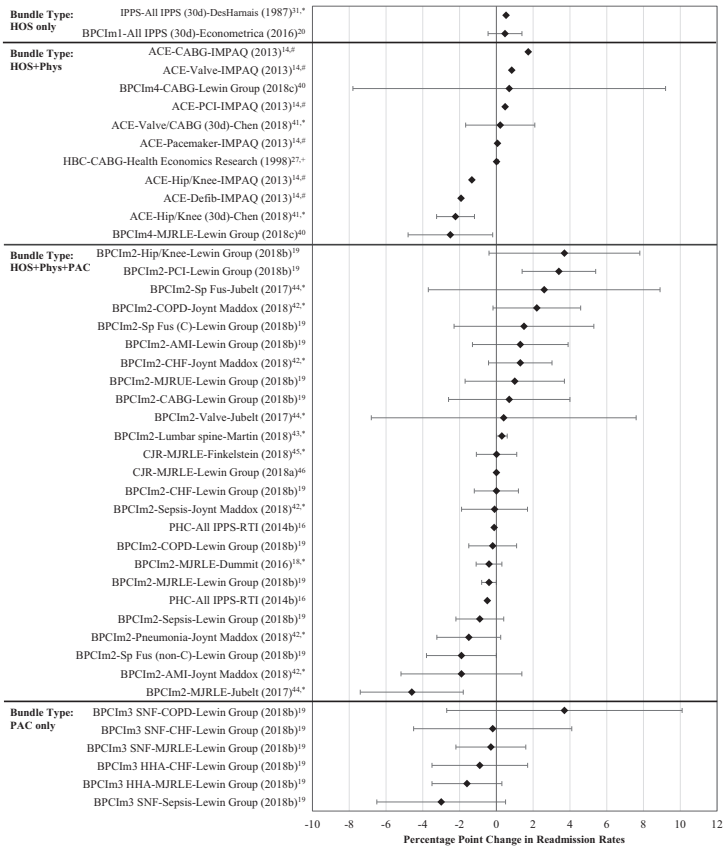
Despite goals and expectations for bundled payment models, few studies have found that initiatives had significant impacts on readmission rates within the episode duration. Out of 44 estimates, only five were significant, three of which showed reductions in rates and two of which showed an increase in rates (Figure 7). Readmission rates significantly decreased for (a) hip and knee replacement episodes under ACE,⁴¹ (b) MJRLE episodes under BPCI Model 4,¹⁹ and (c) MJRLE episodes under BPCI Model 2.⁴⁴ Rates significantly increased for (a) PCI episodes under BPCI Model 2¹⁹ and (b) lumbar spine episodes under BPCI Model 2.⁴³ Among all estimated impacts on readmissions (within 90 days postdischarge), including those that were not significant, 23 estimates were positive and 21 estimates were negative.

Mortality, Emergency Department Use, and Patient Outcomes

Figure 8 shows the estimated effects of bundled payment initiatives on 90-day all-cause mortality rates, unless there was only evidence on 30-day mortality rates. Point estimates include 22 positive and 17 negative estimates across the bundled payment initiative groups. Under IPPS, mortality rates decreased by a substantial amount for patients admitted to the ICU.³⁴ Under BPCI Model 2, mortality rates declined for AMI episodes but increased for COPD episodes.⁴² All other estimates either were not statistically significant or the studies did not provide confidence intervals to evaluate statistical significance.

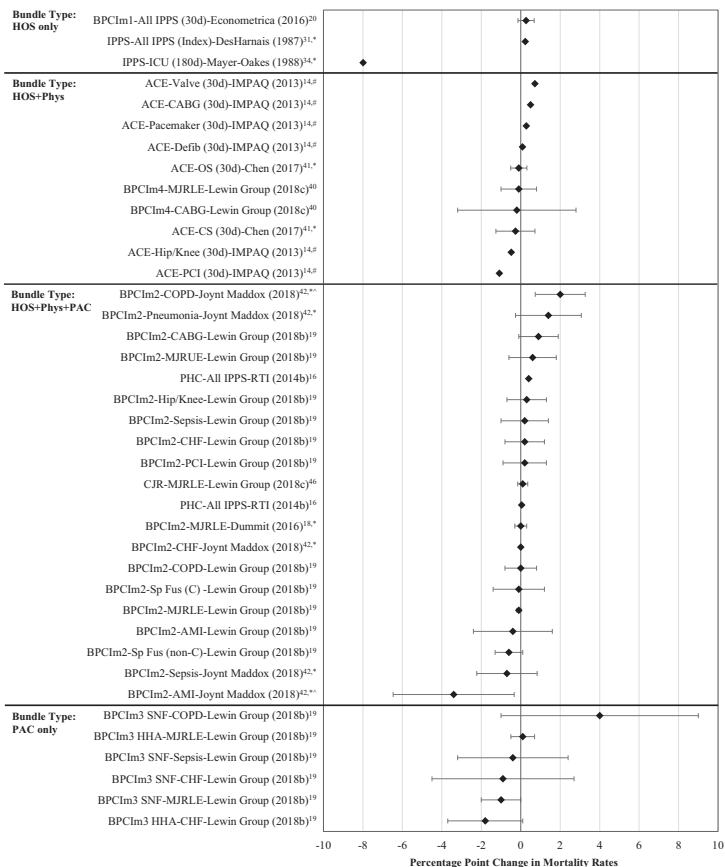
ED admissions within 90 days of discharge were largely unaffected by bundled payments, with one exception (Figure 9). The Lewin Group¹⁹ found a lower likelihood of ED admission for MJRUE episodes under BPCI Model 2. Although most of the findings were not statistically significant, 17 of the 28 (61%) estimates indicated greater likelihood of ED admission with bundled payments.

Figure 7. Bundled Payment Impact on Readmission Rates (90-day rates unless otherwise specified)



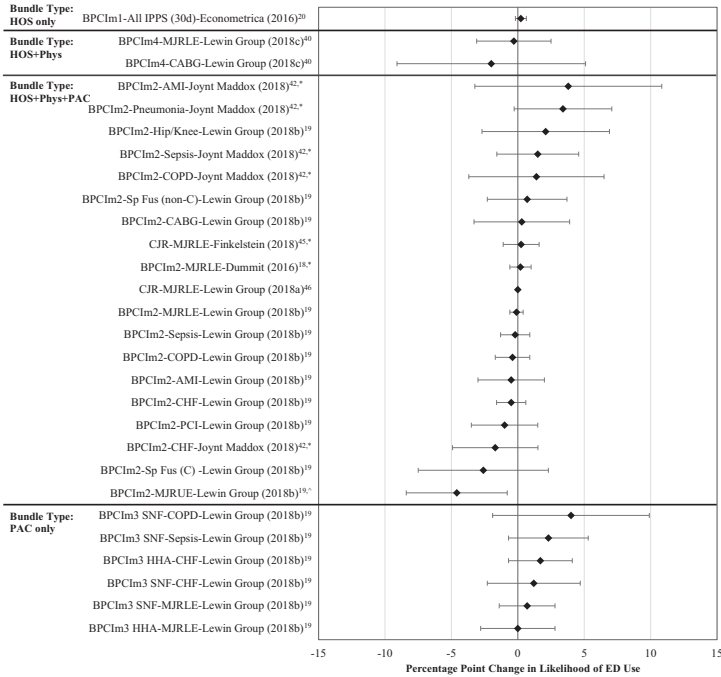
Study estimates and 95% confidence intervals are shown. Most estimates are for 90-day readmission rates. A few studies only provided effects on 30-day rates, which are denoted by (30d) in the label. * indicates peer-reviewed journal publication. ^ indicates estimate was significant at the 5% level. See appendix for list of initiative and condition abbreviations. +HBC-CABG-Health Economics Research (1998)²⁷ provides estimates for each participant; two of the seven participant estimates were statistically significant. For consistency, however, the figure shows the average of the participants' estimates, and the error bars represent the highest and lowest participant estimate. #ACE-IMPAQ (2013)¹⁴ provided odds ratio estimates, which we converted to absolute differences to be comparable to other estimates.

Figure 8. Bundled Payment Impact on Mortality Rates (90-day rates unless otherwise specified)



Study estimates and 95% confidence intervals are shown. * indicates peer-reviewed journal publication. ^ indicates estimate was significant at the 5% level. See appendix for list of initiative and condition abbreviations. Most estimates are for 90-day mortality rates. A few studies provided the effect on rates for other time frames; these are denoted by (30d), (Index), or (180d) in the label. *Index* is the hospital admission that begins the episode of care. #ACE-IMPAQ (2013)¹⁴ provided odds ratio estimates, which we converted to absolute differences to be comparable to other estimates.

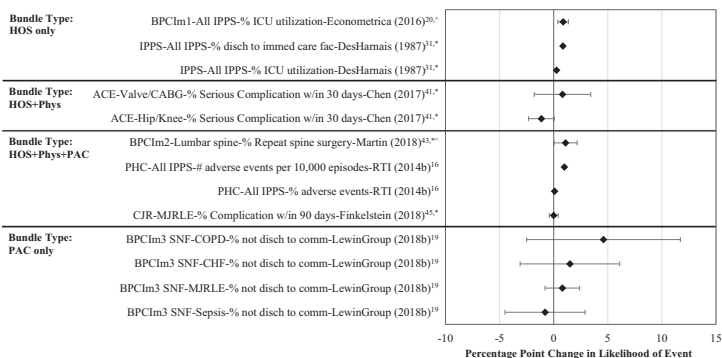
Figure 9. Bundled Payment Impact on Emergency Department Use (90-day rates unless otherwise specified)



Study estimates and 95% confidence intervals are shown. * indicates the study was published in a peer-reviewed journal. ^ indicates that the results were significant at the 5% level. Please refer to the appendix list of abbreviations for initiative and condition abbreviations. Most estimates are for the likelihood of an emergency department admission within 90 days of discharge. One study only provided that within 30 days; this is denoted by (30d) in the label.

Complications and patient outcomes were not measured consistently across studies. The Lewin Group¹⁹ collected survey data on activities of daily living. Other studies analyzed the proportion of episodes with an ICU admission after being discharged,^{20,31} the number of episodes that had a serious complication or adverse event,^{16,41} repeat spine surgery rates,⁴⁵ and the proportions of patients discharged and not discharged

Figure 10a. Bundled Payment Impact on Adverse Patient Outcome Events



Study estimates and 95% confidence intervals are shown. * indicates the study was published in a peer-reviewed journal. ^ indicates that the results were significant at the 5% level. Patient Outcome Abbreviations: *Comm*, community; *Disch*, discharged; *Immed care fac*, immediate care facility. See appendix for list of initiative and condition abbreviations. For BPCI Model 3, the Lewin Group (2018b)¹⁹ provided the change in the percentage of episodes discharged to the community. To keep this consistent with the other measures, which measure adverse patient outcomes, we negated the signs of the estimates. Thus, a larger positive value means a worsening of patient outcomes.

to the community.¹⁹ Figure 10a shows the difference-in-difference estimates for these event-like outcomes. Figure 10b shows the estimates for patient-functioning outcomes, such as improved ambulation after treatment.

Figure 10a shows that bundled payment models were associated with worse event-like outcomes; however, all but two of the findings were not statistically significant. Econometrica²⁰ found that under BPCI Model 1 (which did not include postdischarge care in the bundle), the proportion of patients admitted to the ICU significantly increased relative to the control group. Martin and coauthors⁴³ found that the frequency of repeat surgery significantly increased under BPCI Model 2.

Under IPPS, DesHarnais and colleagues³¹ found an increase in the likelihood of an ICU admission (no statistical significance was provided);

however, Mayer-Oakes and associates³⁴ found that patient functioning after admission to the ICU was better (Figure 10b). Together, the findings from these two studies might suggest that IPPS was associated with a shifting of less severe patients to ICUs.

The Lewin Group^{19,40} reported a mix of positive and negative associations between BPCI Models 2, 3, and 4 and patient functioning, but many of the estimates were not statistically significant. For brevity, we included only a subset of the patient functioning measures the Lewin Group studied in Figure 10b (see Lewin Group reports^{19,40} for more information). For the BPCI models, 39 of 65 estimates (60%) were negative, indicating a worsening of patient functioning. Five estimates were statistically significant, four of which suggested a worsening of outcomes, and one of which suggested an improvement. Patients undergoing MJRLE (nonfracture), hip and knee revisions, or non-cervical spinal fusions at hospitals participating in BPCI Model 2 reported worse patient functioning in terms of self-caring abilities. Under BPCI Model 4, patients undergoing MJRLE who had IRF services reported an improved self-care score. Finkelstein and coauthors⁴⁵ found no change in quality associated with CJR, despite reimbursement being based on quality scores. More research is needed to better evaluate patient functioning outcomes; only three studies in our review analyzed these types of outcomes.

Potential Unintended Consequences: Episode Volume and Case Mix

Bundled payments fix the payment per episode but do not control the number of episodes. Providers could increase profits and success by seeing more patients who are less costly and/or fewer patients who are more costly than average.^{18,50,51} In the way the Medicare initiatives have been designed thus far, bundled prices have adjusted for patient severity only to the extent that 1) the status quo payment system for each provider type adjusted for it and 2) the historical payments reflected changes in patient severity (for retrospective payment initiatives). Some patients may be more costly than other patients within a MS-DRG code and thus providers may triage treatment for these patients, if they know ex-ante they would be more costly. Certain easily observable factors may be associated with higher costs within a MS-DRG; for example, Courtney and

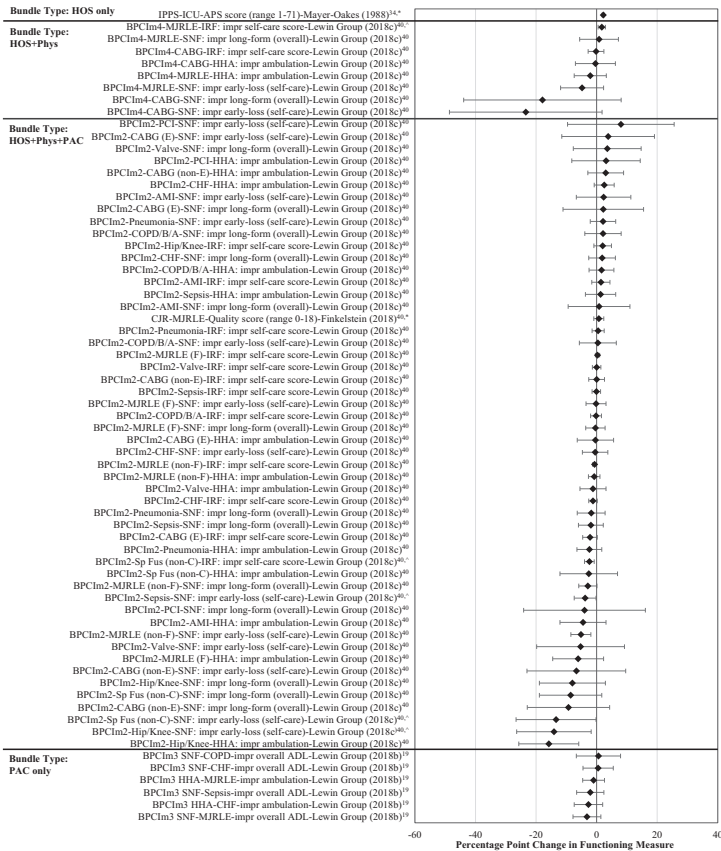
colleagues⁵² found that age, gender, comorbidities, socioeconomic status, and geographic location are associated with different levels of spending for patients under the same episode of a hip or knee replacement.

Episode volume could increase with bundled payment models. For example, if readmissions are not included in the bundle or episode durations are short, hospitals may have an incentive to discharge patients prematurely because a readmission outside the episode duration is considered a new episode. Although, as previously discussed, bundled payment initiatives had mixed effects on readmission rates within the defined episode duration, evidence of increased episode volume could be symptomatic of bundled payments generating more episodes by delaying readmissions.

This possible association is difficult to investigate because of confounding factors. In many early initiatives, CMS gave patient incentives to choose providers who participated in bundled payment contracts with Medicare; this was a way for Medicare to encourage provider participation in the initiatives. For example, in the ACE demonstration, Medicare paid beneficiaries up to 50% of Medicare's cost savings if they went to ACE-participating hospitals for care. Although this promotion was intended to encourage patients to choose ACE-participating hospitals, it may have unintentionally enticed more patients to have the procedures done. According to interviews of managers of hospitals that participated in ACE, a major goal of participation was to increase their hospital volume.^{14,15}

The empirical evidence on bundled payment models' impact on volume is limited. Only five studies in our sample had grade A research designs and reported the statistical significance of their findings. Overall, the evidence in our review suggests that bundled payments increased volume, but many estimates were not significant. IMPAQ¹⁴ found that among ACE participants, the number of defibrillator episodes increased on average by 19.2%, the number of pacemaker episodes decreased by 11.7%, and there was no significant change for other types of episodes (cardiac valve, CABG, PCI, hip/knee replacements, and other cardiac and orthopedic procedures) relative to the control group. Joynt Maddox and coauthors⁴² showed that under BPCI Model 2, the number of episodes per hospital per quarter increased for four out of five conditions they studied, but none of these findings were statistically significant at the 5% level (see Appendix tables S4a–S4e in their article). Martin and associates⁴³ found that lumbar fusion volume for BPCI Model

Figure 10b. Bundled Payment Impact on Patient Functional Status



Study estimates for multiple functioning measures and 95% confidence intervals are shown. * indicates the study was published in a peer-reviewed journal. ^ indicates that the results were significant at the 5% level. See appendix for list of initiative, condition, post-acute care, and patient outcome abbreviations. Mayers-Oakes (1988)³⁴ evaluated the Acute Physiology Score (APS) among patients admitted to an ICU under IPPS. The Lewin Group 3-year evaluation (2018c)⁴⁰ evaluated four measures for BPCI Model 2: improved status or remained completely independent in ambulation/locomotion among patients who received HHA services (HHA: *impr ambulation*); improved status or remained completely independent in long-form ADL function (SNF: *impr long-form (overall)*); improved status or remained completely independent in early-loss ADL function among patients who received SNF services (SNF: *impr early-loss (self-care)*); and improved self-care score among patients who received IRF services (IRF: *impr self-care score*). Two measures are from the Lewin Group 5-year evaluation (2018b)¹⁹ for episodes under BPCI Model 3: improved ambulation, ADL (*impr Ambulation*); and improved overall function, ADL (*impr overall ADL*).

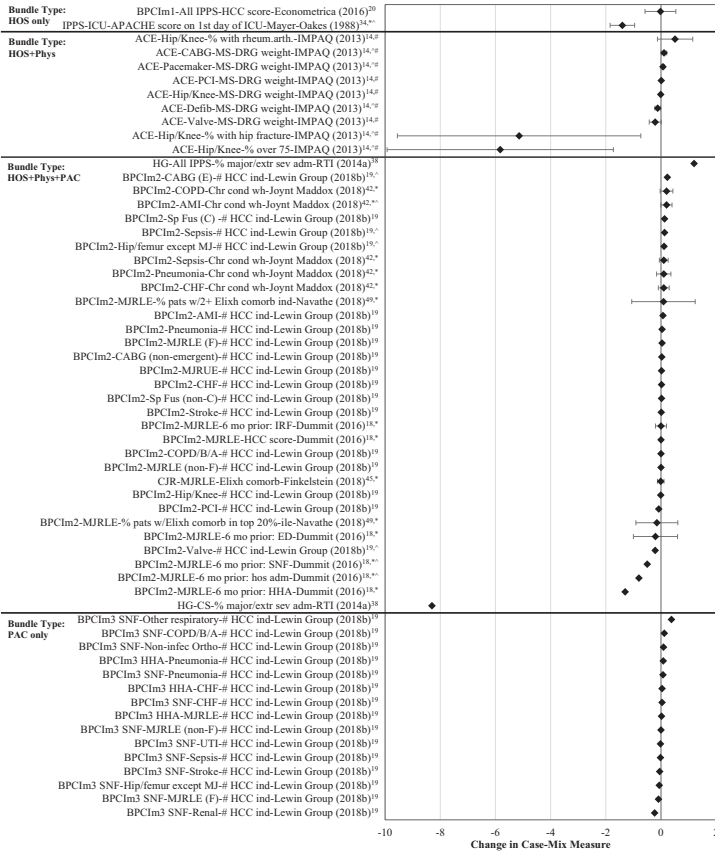
2 risk-bearing hospital participants increased by 5% relative to non-participants' volume, but this finding was not significant ($P = 0.11$). Navathe and colleagues⁴⁹ found that the number of MJRLE episodes per 1,000 Medicare beneficiaries per quarter in markets with BPCI Model 2 participants increased by 0.32% relative to markets without BPCI Model 2 participants, but this difference was not significant ($P = 0.10$). However, they did find a significant 0.39% ($P = 0.04$) increase in the volume of uncomplicated MJRLEs, suggesting a potential provider preference for less-severe patients. Finkelstein and associates⁴⁵ found a nonsignificant 0.05 increase in the number of CJR-eligible MJRLE admissions per 1000 enrollees relative to the control group. RTI^{16,38} also found increases in volume but did not report on the significance of these findings.

Like the evidence on episode volume, the evidence on case-mix selection in our review was limited. Case-mix measures include functions of comorbidity indicators, such as Medicare's hierarchical condition categories (HCCs) and the Elixhauser comorbidity index; MS-DRG weights; the Acute Physiology and Chronic Health Evaluation (APACHE) score (a composite of an acute physiology score, age, and chronic health); SNF or HHA use prior to episode; and measures of age.

Many of the estimated effects on case-mix selection were not significant at the 5% level (Figure 11). Of the 59 estimates in our sample, 23 (39%) indicated a change toward treating less-severe patients under bundled payments. Among the 13 statistically significant results, seven (54%) suggested a shift toward selecting less-severe patients. APACHE score data showed that the severity of inpatient episodes in which patients were admitted to the ICU decreased under IPPS.³⁴ IMPAQ reported that, relative to the control group, fewer patients with hip and knee replacement episodes under ACE had a hip fracture, and fewer were older than 75 years.¹⁴ In the same report, the MS-DRG weight was lower (indicating lower severity) for defibrillator episodes under ACE but higher for CABG and pacemaker episodes.¹⁴

According to Dummit and coauthors (the investigators for the Lewin Group), the severity of MJRLE cases, as measured by whether the patient had HHA or SNF services within 6 months prior to the episode, lessened under BPCI Model 2.¹⁸ Although not shown in Figure 11, Dummit and associates¹⁸ also found that, compared to nonparticipants, hospitals participating in BPCI Model 2 had greater reductions in both the number of Medicaid-eligible patients and the number of patients

Figure 11. Bundled Payment Impact on Case Mix



Study estimates and 95% confidence intervals are shown. * indicates the study was published in a peer-reviewed journal. ^ indicates that the results were significant at the 5% level. The Change in Case Mix Measure is the difference-in-differences estimate, meaning the change in, for example, the MS-DRG weight observed in the initiative group of providers relative to the change observed in a control group of providers. #ACE-IMPAQ (2013)¹⁴ provided odds ratio estimates, which we converted to absolute differences to be comparable to other estimates. See appendix for list of initiative and condition abbreviations. Case mix abbreviations: *APACHE* score, physiological functioning score; *Chr cond wh*, Chronic condition warehouse, Medicare-supplied comorbidity measure; *Elixh comorb*, Elixhauser comorbidity index score; *HCC*, Medicare's hierarchical condition category; *ICU*, intensive care unit; *Ind*, indicator; *MS-DRG*, Medicare's Medical Severity Diagnosis Related Group; # *HCC ind*, number of patient episodes flagged with an HCC indicator; % *major/extr sev adm*, % of admissions of major or extreme severity.

who had an inpatient admission prior to the index hospitalization. In the Lewin Group 5-year evaluation report,¹⁹ the investigators found significant increases in the number of patients with HCC indicators in sepsis, AMI, and hip and femur (except major joint) episodes; however, they found a significant decrease in the number for cardiac valve episodes.

Discussion

Policy Implications

CMS has experimented with various features in designing their bundled payment initiatives, each with an expected outcome. We draw a few conclusions from the literature about these experiments. First, as designs have evolved, providers have changed their focus on which areas to cut costs. Under IPPS, hospitals reduced length of stay substantially, but the change in length of stay was much smaller under most of the more recent initiatives, with the single exception of BPCI Model 2.⁴⁴ In the several initiatives following IPPS that introduced gainsharing with physicians, hospitals focused on reducing internal production costs. Under more recent initiatives that had more inclusive bundle definitions and episode durations, such as BPCI Models 2 and 3 and CJR, the focus turned to savings in post-acute care settings.

Second, one of the driving factors in reducing Medicare spending has been the size of the discount on the price for bundled services. When discounts were larger than 5%, as in HBC and ACE, Medicare saved more. Such savings might be explained by changes in patient case-mix selection. When discounts were less than 5%, as in the BPCI models and CJR, savings were lower.

Third, as the bundle definition became more inclusive, there was less shifting of services to outside the bundle. Outpatient services increased if they were not included in the bundle (as shown under IPPS). Post-acute care services also increased when they were not included in the bundle. Three of the five initiatives promulgated by the ACA have included post-acute care services in the bundle, and studies found substantial reductions in SNF and IRF services associated with these models.

Fourth, the provider type at greatest financial risk seems to have affected whether spending per episode was reduced. The largest downward effect on Medicare spending was found under BPCI Model 3, in which

post-acute care providers were at greatest financial risk. When SNFs were the episode awardees in BPCI Model 3 contracts, Medicare payments to SNFs for certain conditions (MJRLE, CHF, sepsis) decreased by \$1,500-\$2,000 per episode (9%-14%), and overall spending per episode decreased by 3%-6%. The reductions seem to have been achieved by reducing SNF services and shifting to HHA services.

Furthermore, reductions in Medicare per-episode payments to SNFs under BPCI Model 3 were larger than they were for BPCI Model 2, in which a hospital or physician provider group was the episode awardee. Under both BPCI Model 2 and BPCI Model 3, SNFs were paid per diem rates, and episode payments were reconciled ex post with the target prices for the episodes. However, financial risk was not distributed in the same way in the two models: Under BPCI Model 3, SNF awardees bore the entire financial risk for the episode; under BPCI Model 2, SNF incentives may or may not have been aligned with the hospital or physician group awardee, depending on the details of the gainsharing arrangement and whether there were contractual relationships between the hospital or physician group awardee and the SNF prior to participating in BPCI Model 2.

Based on our discussions with a few hospital participants, some BPCI Model 2 hospital awardees (which had limited time to set up contracts with other providers) chose to focus their coordination efforts (and gainsharing) with physician groups instead of post-acute care providers. According to a 2016 *New England Journal of Medicine* Catalyst Report, approximately 40% of SNFs are not contractually integrated with other providers.⁵³ Another reason that per-episode payments to SNFs may not have decreased by as much under BPCI Model 2 as they did under Model 3 is that the composition of SNF patients changed in different ways under the two models. Hospital participant awardees of BPCI Model 2 achieved cost savings by shifting patients from higher-cost IRFs to lower-cost SNFs.⁵⁴ If IRF patients were in need of more post-acute care services, this shift might reduce overall post-acute care spending but would probably not lower SNF spending. Furthermore, there might be shifts of relatively healthier patients from SNFs to HHAs.

Fifth, with few exceptions, bundled payment initiatives were not associated with significant changes in mortality, ED use, or adverse events. Among the few significant changes in readmission rates, some initiative-conditions increased and some decreased, with no relation to the bundle definition. The findings on patient functioning indicated that more than

half of the estimates corresponded to worsening of function under BPCI Models 2, 3, and 4; however, only a few types of episodes (under BPCI Model 2) were significant. It will be important to continue studying CJR because the program is not voluntary, and it will be the first to make provider reimbursement contingent on quality scores. If the canceled initiatives resume, they will also warrant investigation.

Finally, bundled payment initiatives seem to have been successful in lowering spending for certain medical conditions, such as MJRLEs. Some commentators conjecture that MJRLE (compared to infectious diseases or sepsis) is an anticipated, elective procedure, for which providers may have time to more reliably reduce costs. Kivlahan and coauthors⁵⁴ found that Medicare spending on post-acute care services for certain BPCI participants was on average 67% of the total spending for MJRLE; by comparison, post-acute care spending on COPD, CHF, CABG, and cardiac valve surgery patients under the same bundled payment initiative was between 28% and 47%.

Some policymakers and researchers postulate that there may be a trade-off between using post-acute care services and readmission and ED use rates. We found mixed results. For example, hip and knee replacement episodes under BPCI Model 2 were associated with the largest reductions in post-acute care services (SNF, IRF, or LTCH). They also were associated with the largest increase in 90-day readmission rates and one of the largest increases in ED use; however, the estimates on readmission rates and ED use were not significant. MJRLE episodes under BPCI Model 2 were also associated with reductions in the use of post-acute care services. However, they were associated with small, not significant reductions in readmission rates and ED use. Jubelt and colleagues⁵⁵ found that for one hospital participant in BPCI Model 2, the shift from institutional post-acute care to home health services was not associated with a change in readmission rate.

Limitations

Our review has a few limitations. Any review can only draw on published studies, which could be subject to publication bias. Furthermore, many provider groups have attempted to form bundled payment contracts but failed to agree on the terms; no results from such attempts have been published.

Another limitation is that only selected medical conditions and interventions have been the focus of bundled payment initiatives and their evaluations. Those conditions and interventions may have been selected because they were most costly or had high rates of growth. Moreover, in many cases, participating organizations could select the conditions to be bundled and may have chosen only those conditions for which they anticipated spending less than the target price. Furthermore, researchers may tend to focus on conditions that have sufficient sample size or areas in which bundled payments may have been effective.

Conclusion

Looking to the future, for Medicare to achieve savings through higher discounts or more participants, providers need to see the potential to reduce internal production costs. Providers have reduced production costs per episode by targeting post-acute care services, prices for medical devices and supplies, and hospital processes. However, these options may not be possible for all providers. It is important to understand not only the impact of various bundled payment designs but also the reasons for provider participation in some models but not in other models/conditions.

We can, of course, only study bundled payment contracts that have been formed successfully, in which the provider and payer agree to the terms—and there have been many both within and outside of Medicare.^{56,57} However, not all efforts to form such contracts have been successful.⁵⁸ Moreover, some that have been initiated have had attrition issues; for example, this has been a concern for some contracts under BPCI and CJR models (Table 1). Payers and provider groups have found it difficult to agree on how much risk should be borne by providers. Payers would like a broader inclusion of services and longer episodes of care, whereas providers would like to narrow inclusion and limit episode lengths. In addition, payers are worried that bundled payments incentivize providers to turn away sicker and more complex patients,^{58,59} and we have found some evidence that this concern has merit. Literature reviews by Steenhuis and colleagues⁶⁰ and the MITRE Corporation⁶¹ provide considerations for parties involved in procuring a bundled payment contract.

Medicare's experimentation with bundled payments has had positive externalities. Providers can take advantage of economies of scale from information technology systems and processes. Private insurers can learn from initiative results as they test their own bundled payment models. There is room for continued, gradual expansion and refinement of bundled payment initiatives. Ongoing adjustments to the design features and monitoring patient outcomes are critical to the success or failure of bundled payments.

References

1. McLawhorn AS, Buller LT. Bundled payments in total joint replacement: keeping our care affordable and high in quality. *Curr Rev Musculoskelet Med.* 2017;10:370-377. <https://doi.org/10.1007/s12178-017-9423-6>.
2. Shih T, Chen LM, Nallamotheu BK. Will bundled payments change health care? Examining the Evidence Thus Far in Cardiovascular Care. *Circulation.* 2015;131(1): 2151-2158. <https://doi.org/10.1161/circulationaha.114.010393>
3. Siddiqi A, White PB, Mistry JB, et al. Effect of bundled payments and health care reform as alternative payment models in total joint arthroplasty: a clinical review. *J Arthroplasty.* 2017;32(8):2590-2597. <https://doi.org/10.1016/j.arth.2017.03.027>.
4. Hung M, Nielson D, Hansen A, Gu Y, Bounsanga J, Voss M. The effects of bundled payment in health care: a comprehensive review of the literature. *J Health Human Services Admin.* 2018;41(2):31.
5. Hussey PS, Mulcahy AW, Schnyer C, Schneider EC. Bundled payment: effects on health care spending and quality. Closing the quality gap: revisiting the state of the science. Agency for Healthcare Research and Quality publication no. 12-E007-EF. 2012. https://effectivehealthcare.ahrq.gov/sites/default/files/pdf/bundled-payments-quality-effects_research.pdf. Accessed June 29, 2017.
6. Davis K, Stremikis K, Squires D, Schoen C. 2014. How the U.S. Health Care System Compares Internationally. <https://www.commonwealthfund.org/publications/fund-reports/2014/jun/mirror-mirror-wall-2014-update-how-us-health-care-system>. Accessed June 29, 2020.
7. The US Burden of Disease Collaborators. The State of US Health, 1990–2016: Burden of Diseases, Injuries, and Risk Factors Among US States. *JAMA.* 2018;319(14):1444-1472. <https://doi.org/10.1001/jama.2018.0158>

8. Baicker K, Chernerw M. Alternative alternative payment models. *JAMA Intern Med.* 2017;177(2):222-223. <https://doi.org/10.1001/jamainternmed.2016.8280>
9. Lave JR. The effect of the Medicare prospective payment system. *Annu Rev Public Health.* 1989;10(1):141-161. <https://doi.org/10.1146/annurev.pu.10.050189.001041>.
10. Altman SH, Eichenholz J. "Inflation in the Health Industry — Causes and Cures." In *Health: A Victim or Cause of Inflation* (Michael Zubkoff, ed.). New York: PRODIST (published for Milbank Memorial Fund), 1976, pp. 7-30.
11. Chulis GS. Assessing Medicare's prospective payment system for hospitals. *Med Care Res Rev.* 1989;48(2):167-206.
12. Feinglass J, Holloway JJ. The initial impact of the Medicare prospective payment system on U.S. health care: a review of the literature. *Med Care Rev.* 1991;48(1):91-115.
13. Navathe A, Troxel AB, Liao JM, et al. Cost of joint replacement using bundled payment models. *JAMA Intern Med.* 2017;177(2):214-222. <https://doi.org/10.1001/jamainternmed.2016.8263>.
14. IMPAQ International; Urdapilleta O, Weinberg D, Pedersen S, Kim G, Cannon-Jones S, Woodward J. Evaluation of the Medicare Acute Care Episode (ACE) demonstration: final evaluation report. <https://downloads.cms.gov/files/cmami/ACE-EvaluationReport-Final-5-2-14.pdf>. Published 2013. Accessed March 6, 2017.
15. Vesely R. An ACE in the deck? Bundled-payment demo shows returns for hospitals, physicians, patients. *Modern Healthcare.* February 7, 2011. <http://www.modernhealthcare.com/article/20110207/MAGAZINE/110209990>. Accessed January 21, 2017.
16. RTI International; Greenwald L, Adamache W, Cole-Beebe M, Amico P, Hunter E, Baker B. Evaluation of the Medicare physician hospital collaboration demonstration: final report. https://downloads.cms.gov/files/cmami/PHC_FINAL-RPT_September2014.pdf. Published 2014. Accessed January 26, 2017.
17. Centers for Medicare and Medicaid Services. Medicare Acute Care Episode Demonstration for orthopedic and cardiovascular surgery. <https://innovation.cms.gov/Files/x/ACE-Web-Page.pdf>. Accessed October 17, 2016.
18. 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(12):1267-1278. <https://doi.org/10.1001/jama.2016.12717>.

19. Lewin Group; Dummit L, Marrufo G, Marshall J, et al. CMS Bundled Payments for Care Improvement Initiative Models 2–4: Year 5 Evaluation & Monitoring Annual Report. <https://downloads.cms.gov/files/cmimi/bpci-models2-4-yr5evalrpt.pdf>. Published October 2018. Accessed February 28, 2019.
20. Econometrica, Inc; Sheppard M, et al. Evaluation and monitoring of the Bundled Payments for Care Improvement Model 1 Initiative: Annual Report 2015. <https://innovation.cms.gov/files/reports/bpci-mdl1yr2annrpt.pdf>. Published May 2016. Accessed February 28, 2019.
21. Levy S, Bagley N, Rajkumar R. Reform at risk—mandating participation in alternative payment plans. *N Engl J Med*. 2018;378(18):1663-1665. <https://doi.org/10.1056/NEJMp1800483>.
22. Centers for Medicare and Medicaid Services. Comprehensive Care for Joint Replacement Model. <https://innovation.cms.gov/initiatives/cjr>. Accessed May 17, 2019.
23. Abt Associates; Reardon LB, Mason C. *Medicare Cataract Surgery Alternate Payment Demonstration: Final Evaluation Report*. <https://archive.org/details/medicarecataract00rear/mode/2up>. Published 1997. Accessed June 2, 2017.
24. Coulam RF, Gaumer GL. Medicare's prospective payment system: a critical appraisal. *Health Care Financ Rev*. 1992;1991(Suppl):45-77.
25. Menke, T. (1990). Impacts of PPS on Medicare Part B expenditures and utilization for hospital episodes of care. *Inquiry* 27(2): 114-126.
26. Cromwell J, Dayhoff DA, Thoumaian AH. Cost savings and physician responses to global bundled payments for Medicare heart bypass surgery. *Health Care Financ Rev*. 1997;19(1): 41-57.
27. Health Economics Research; Cromwell J, Dayhoff DA, McCall NT, et al. Medicare Participating Heart Bypass Center Demonstration. Final Report Volumes 1 and 2. <https://innovation.cms.gov/medicare-demonstrations/medicare-participating-heart-bypass-center-demonstration>. Published 1998. Accessed June 2, 2017.
28. Lave JR. The impact of the Medicare prospective payment system and recommendations for change. *Yale J Regulation*. 1990;7:499-528.
29. Rosenberg MA, Browne MJ. The impact of the Inpatient Prospective Payment System and diagnosis-related groups: a survey of the literature. *North Am Actuarial J*. 2001;5(4): 84-94. <https://doi.org/10.1080/10920277.2001.10596020>.

30. Feder J, Hadley J, Zuckerman S. How did Medicare's Prospective Payment System affect hospitals? *N Engl J Med.* 1987;317(14):867-873.
31. Desharnais S, Kobrinski E, Chesney J, Long M, Ament R, Fleming S. The Early Effects of the Prospective Payment System on Inpatient Utilization and the Quality of Care. *Inquiry* 1987;24(1): 7-16.
32. Guterman S, Eggers PW, Riley G, Greene TF, Terrell SA. The first 3 years of Medicare prospective payment: an overview. *Health Care Financ Rev.* 1988;9(3):67-77.
33. Sloan F, Morrissey M, Valvona J. Effects of the Medicare Prospective Payment System on hospital cost containment: an early appraisal. *Milbank Q.* 1988;66(2):191-220. <http://www.jstor.org/stable/3350030>.
34. Mayer-Oakes A, Oye R, Leake B, Brook R. The early effect of Medicare's Prospective Payment System on the use of medical intensive care services in three community hospitals. *JAMA.* 1988;260(2):3146-3149.
35. Frank R, Lave J, Taube C, Rupp A, Goldman H. The impact of Medicare's Prospective Payment System on psychiatric patients treated in scatterbeds. National Bureau of Economic Research (NBER) working paper no. 2030. September 1986. <https://www.nber.org/papers/w2030>. Accessed September 3, 2017.
36. Glickman A, Dinh C, Navathe AS. The current state of evidence on bundled payments. *LDI Issue Brief.* 2018;18(22). <https://ldi.upenn.edu/brief/current-state-evidence-bundled-payments>. Accessed May 17, 2019.
37. Piccinin MA, Sayeed Z, Kozlowski R, Bobba V, Knesek D, Frush T. Bundle payment for musculoskeletal care. *Orthop Clin North Am.* 2018;49(2):147-156. <https://doi.org/10.1016/j.ocl.2017.11.003>.
38. RTI International; Greenwald L, Adamache W, Cole-Beebe M, et al. Evaluation of the Medicare gainsharing demonstration, final report. <https://innovation.cms.gov/files/reports/medicaregainsharingappendix.pdf>. Published January 2014. Accessed September 9, 2017.
39. Liu CF, Subramanian S, Cromwell J. Impact of Global Bundled Payments on Hospital Costs of Coronary Artery Bypass Grafting. *J Health Care Financ* 2001;27(4):39-54.
40. Lewin Group; Dummit L, Marrufo G, Marshall J, et al. CMS bundled payments for care improvement initiative models 2-4: Year 3 Evaluation & Monitoring Annual Report. <https://downloads.cms.gov/files/cmimi/bpci-models2-4yr3evalrpt.pdf>. Revised 2018. Accessed March 28, 2019.

41. Chen LM, Ryan AM, Shih T, Thumma JR, Dimick JB. Medicare's Acute Care Episode Demonstration: Effects of bundled payments on costs and quality of surgical care. *Health Serv Res.* 2018;53(2):632-648. <https://doi.org/10.1111/1475-6773.12681>.
42. Joynt Maddox KE, Orav EJ, Zheng J, Epstein AM. Evaluation of Medicare's bundled payments initiative for medical conditions. *N Engl J Med.* 2018;379(3):260-269. <https://doi.org/10.1056/NEJMsa1801569>.
43. Martin B, Lurie J, Farrokhi F, McGuire K, Mirza SK. Early effects of Medicare's bundled payment for care improvement program for lumbar fusion. *Spine (Phila Pa 1976).* 2018;43(10):705-711. <https://doi.org/10.1097/BRS.0000000000002404>
44. Jubelt LE, Goldfeld KS, Blecker SB, et al. Early lessons on bundled payment at an academic medical center. *J Am Acad Orthop Surg* 2017;25: 654-663. <https://doi.org/10.5435/JAAOS-D-16-00626>.
45. Finkelstein A, Ji Y, Mahoney N, Skinner J. Mandatory Medicare bundled payment program for lower extremity joint replacement and discharge to institutional postacute care: interim analysis of the first year of a 5-year randomized trial. *JAMA.* 2018;320(9):892-900. <https://doi.org/10.1001/jama.2018.12346>.
46. Lewin Group; Dummit L, Marrufo G, Marshall J, et al. CMS Comprehensive Care for Joint Replacement Model: Performance Year 1 Evaluation Report. <https://innovation.cms.gov/files/reports/cjr-firstannrpt.pdf>. Published August 2018. Accessed May 9, 2020.
47. Fitzgerald JF, Fagun LF, Tierney WN, Dittus RS. Changing pattern of hip fracture care before and after implementation of the prospective payment system. *JAMA.* 1987;258:218-221.
48. Medicare Payment Advisory Commission. March 2017 Report to Congress: Medicare Payment Policy. <http://www.medpac.gov/documents/-reports>. Accessed November 6, 2019.
49. Navathe AS, Liao JM, et al. Association of hospital participation in a Medicare bundled payment program with volume and case mix of lower extremity joint replacement episodes. *JAMA.* 2018;320(9):901-910. <https://doi.org/10.1001/jama.2018.12345>. Published correction appears in *JAMA.* 2018;320(22):2381. <https://doi.org/10.1001/jama.2018.17985>.
50. Crosson FJ; Medicare Payment Advisory Commission. MedPAC comment on CMS's proposed rule on the Comprehensive Care for Joint Replacement payment model for acute care hospitals.

- CMS-5516-P. August 15, 2015. <http://www.medpac.gov/docs/default-source/comment-letters/medpac-comment-on-cms-s-proposed-rule-on-the-comprehensive-care-for-joint-replacement-payment-model-.pdf?sfvrsn=0>. Accessed March 29, 2020.
51. Mechanic R. *Issue Brief. Medicare's Bundled Payment Initiatives: Considerations for Providers*. Washington, DC: American Hospital Association; 2016.
 52. Courtney PM, Bohl DD, Lau EC, Ong KL, Jacobs JJ, Della Valle CJ. Risk adjustment is necessary in Medicare bundled payment models for total hip and knee arthroplasty. *J Arthroplasty*. 2018;33(8):2368-2375. <https://doi.org/10.1016/j.arth.2018.02.095>.
 53. Compton-Phillips A, Mohta NS. Care Redesign Survey: Strengthening the Post-Acute Care Connection. *New England Journal of Medicine Catalyst*. November 10, 2016. <https://catalyst.nejm.org/strengthening-post-acute-care-connection/>. Accessed November 4, 2019.
 54. Kivlahan C, Orłowski JM, Pearce J, Walradt J, Baker M, Kirch M. Taking risk: early results from teaching hospitals' participation in the Center for Medicare and Medicaid Innovation Bundled Payments for Care Improvement Initiative. *Acad Med*. 2016;91(7):936-942. <https://doi.org/10.1097/ACM.0000000000001121>.
 55. Jubelt L, Goldfeld K, Chung W. Changes in discharge location and readmission rates under Medicare bundled payment. *JAMA Intern Med*. 2016;176(1): 115-117.
 56. Slotkin JR, Ross OA, Coleman MR, Ryu J. Why GE, Boeing, Lowe's, and Walmart are directly buying health care for employees. *Harvard Business Review*. June 8, 2017. <https://hbr.org/2017/06/why-ge-boeing-lowes-and-walmart-are-directly-buying-health-care-for-employees>. Accessed February 1, 2020.
 57. Matchar D, Nguyen H, Tian Y. Bundled payment and care of acute stroke. *Stroke*. 2015;46(5):1414-1421. <https://doi.org/10.1161/STROKEAHA.115.009089>.
 58. Ridgely MS, De Vries D, Bozic KJ, Hussey PS. Bundled payment fails to gain a foothold in California: the experience of the IHA bundled payment demonstration. *Health Affairs*. 2014;33(8): 1345-1352.
 59. Hussey PS, Ridgely MS, Rosenthal MB. The PROMETHEUS bundled payment experiment: slow start shows problems in implementing new payment models. *Health Affairs*. 2011;30(11):2116-2124. <https://doi.org/10.1377/hlthaff.2011.0784>.

60. Steenhuis S, Struijs J, Koolman X, Ket J, Hijden EVD. Unraveling the complexity in the design and implementation of bundled payments: a scoping review of key elements from a payer's perspective. *Milbank Q.* 2020;98(1):197-222. <https://doi.org/10.1111/1468-0009.12438>.
61. MITRE Corporation. Contracting for bundled payment. Prepared for the Centers for Medicare and Medicaid Services. 2012. https://www.mitre.org/sites/default/files/pdf/Contracting_Bundled_Payment.pdf. Accessed October 26, 2016.

Funding/Support: This work was supported by Partnered Evidence-Based Policy Resource Center's Quality Enhancement Research Initiative and Health Services Research and Development grants PEC 16-001 and SDR 16-196, respectively, from the US Department of Veterans Affairs. The views expressed in this article are those of the authors and do not necessarily represent the views of the US Department of Veterans Affairs, the US Government, University of Maryland Baltimore County, Boston University, Harvard University, or any other organization with which the authors are affiliated.

Conflict of Interest Disclosures: None

Acknowledgments: We are extremely grateful for the research assistance by Megha Parikh and Elsa Pearson and for our funding sources. This work was supported by the Quality Enhancement Research Initiative and Health Services Research and Development of the Veterans Health Administration.

Address correspondence to: Christine A. Yee, PhD, Department of Economics, University of Maryland Baltimore County, 1000 Hilltop Circle, Baltimore, MD 21250 (email: christine.yee@gmail.com).

Appendix

List of Abbreviations

Bundled payment initiative abbreviations

ACE	Medicare Acute Care Episode Demonstration
AMI	Model Acute Myocardial Infarction Model
BPCIm1	Bundled Payments for Care Improvement Model 1
BPCIm2	Bundled Payments for Care Improvement Model 2
BPCIm3	Bundled Payments for Care Improvement Model 3
BPCIm4	Bundled Payments for Care Improvement Model 4
CABG	Model Coronary Artery Bypass Graft Model
CJR	Comprehensive Care for Joint Replacement Model Initiative
CR	Model Cardiac Rehabilitation Incentive Payment Model

CSAP	Cataract Surgery Alternate Payment Demonstration
HBC	Medicare Participating Heart Bypass Center Demonstration
HG	Medicare Hospital Gainsharing Demonstration
IPPS	Inpatient Prospective Payment System
PHC	Physician Hospital Collaboration Demonstration
SHFFT	Model Surgical Hip and Femur Fracture Treatment Model

Post-acute care abbreviations

HH	home health
HHA	home health agency
Inst PAC PAC	post-acute care services in an institutional setting (SNF, IRF, or LTCH) given that a patient received post-acute care services
IRF	institutional rehabilitation facility
LTCH	long-term care hospital
PAC	any or total post-acute care services
SNF	skilled nursing facility

Conditions/outcomes abbreviations

ADL	activities of daily living
AMI	acute myocardial infarction
CABG	coronary artery bypass graft
CABG (E)	coronary artery bypass graft, emergent
CABG (non-E)	coronary artery bypass graft, non-emergent
CHF	congestive heart failure
COPD	chronic obstructive pulmonary disease
COPD/B/A	chronic obstructive pulmonary disease, bronchitis, asthma
CS	cardiac surgery
Defib	defibrillator surgery
Hip/Knee	hip or knee revision
MJ	major joint replacement
MJRLE	major joint replacement of the lower extremity
MJRLE (F)	major joint replacement of the lower extremity, fracture
MJRLE (non-F)	major joint replacement of the lower extremity, non-fracture
MJRUE	major joint replacement of the upper extremity
OS	orthopedic surgery

PCI	percutaneous coronary intervention
Sp Fus (C)	spinal fusion, cervical
SP Fus (non-C)	spinal fusion, non-cervical
UTI	urinary tract infection
Valve	cardiac valve surgery

Patient characteristics/outcomes abbreviations

ADL	activities of daily living
DRG	diagnosis related group
ICU	intensive care unit
MS-DRG	Medicare severity diagnosis related group
OR	operating room

Research design abbreviations

Pre-Post	compare pre- and post-implementation outcome measures among bundled payment initiative participants
DiD	difference-in-differences
LATE	local average treatment effect

Other abbreviations

ASC	ambulatory surgical center
CMS	Centers for Medicare and Medicaid Services
CMMI	Center for Medicare and Medicaid Innovation

Supplementary Material

Additional supporting information may be found in the online version of this article at <http://onlinelibrary.wiley.com/journal/10.1111/ISSN1468-0009>:

Appendix Figure 1 (Pairing with Figure 3). Bundled Payment Impact on Medicare Spending per Episode (Percent Change)

Appendix Figure 2 (Pairing with Figure 4). Bundled Payment Impact on Length of Stay (Percent Change)

Appendix Figure 3 (Pairing with Figure 5). Bundled Payment Impact on Likelihood of Post-Acute Care Utilization (Percent Change)

Appendix Figure 4a (Pairing with Figure 6a). Bundled Payment Impact on Total Post-Acute Care Spending (Percent Change)

Appendix Figure 4b (Pairing with Figure 6b). Bundled Payment Impact on Post-Acute Care Spending for Institutional Rehabilitation Facility Services (Percent Change)

Appendix Figure 4c (Pairing with Figure 6c). Bundled Payment Impact on Post-Acute Care Spending on Home Health Agency Services (Percent Change)

Appendix Figure 4d (Pairing with Figure 6d). Bundled Payment Impact on Post-Acute Care Spending for Skilled Nursing Facility Services (Percent Change)

Appendix Figure 5 (Pairing with Figure 7). Bundled Payment Impact on Readmission Rates, 90-day unless otherwise specified (Percent Change)

Appendix Figure 6 (Pairing with Figure 8). Bundled Payment Impact on Mortality Rates, 90-day unless otherwise specified (Percent Change)

Appendix Figure 7 (Pairing with Figure 9). Bundled Payment Impact on Emergency Department Use, 90-day unless otherwise specified (Percent Change)

Appendix Figure 8a (Pairing with Figure 10a). Bundled Payment Impact on Adverse Patient Outcome Events (Percent Change)

Appendix Figure 8b (Pairing with Figure 10b). Bundled Payment Impact on Patient Functional Status (Percent Change)

Appendix Figure 9 (Pairing with Figure 11). Bundled Payment Impact on Case Mix (Percent Change)