2019

Two-year evaluation of mandatory bundled payments for joint replacement

Michael L. Barnett
Harvard University

Andrew Wilcock
Harvard University

J. Michael McWilliams
Harvard University

Arnold M. Epstein
Harvard University

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

See next page for additional authors

SPECIAL ARTICLE

Two-Year Evaluation of Mandatory Bundled Payments for Joint Replacement

Michael L. Barnett, M.D., Andrew Wilcock, Ph.D., J. Michael McWilliams, M.D., Ph.D., Arnold M. Epstein, M.D., Karen E. Joynt Maddox, M.D., M.P.H., E. John Orav, Ph.D., David C. Grabowski, Ph.D., and Ateev Mehrotra, M.D., M.P.H.

ABSTRACT

BACKGROUND

In 2016, Medicare implemented Comprehensive Care for Joint Replacement (CJR), a national mandatory bundled-payment model for hip or knee replacement in randomly selected metropolitan statistical areas. Hospitals in such areas receive bonuses or pay penalties based on Medicare spending per hip- or knee-replacement episode (defined as the hospitalization plus 90 days after discharge).

METHODS

We conducted difference-in-differences analyses using Medicare claims from 2015 through 2017, encompassing the first 2 years of bundled payments in the CJR program. We evaluated hip- or knee-replacement episodes in 75 metropolitan statistical areas randomly assigned to mandatory participation in the CJR program (bundled-payment metropolitan statistical areas, hereafter referred to as "treatment" areas) as compared with those in 121 control areas, before and after implementation of the CJR model. The primary outcomes were institutional spending per hip- or knee-replacement episode (i.e., Medicare payments to institutions, primarily to hospitals and post–acute care facilities), rates of postsurgical complications, and the percentage of "high-risk" patients (i.e., patients for whom there was an elevated risk of spending — a measure of patient selection). Analyses were adjusted for the hospital and characteristics of the patients and procedures.

RESULTS

From 2015 through 2017, there were 280,161 hip- or knee-replacement procedures in 803 hospitals in treatment areas and 377,278 procedures in 962 hospitals in control areas. After the initiation of the CJR model, there were greater decreases in institutional spending per joint-replacement episode in treatment areas than in control areas (differential change [i.e., the between-group difference in the change from the period before the CJR model], −$812, or a −3.1% differential decrease relative to the treatment-group baseline; P<0.001). The differential reduction was driven largely by a 5.9% relative decrease in the percentage of episodes in which patients were discharged to post–acute care facilities. The CJR program did not have a significant differential effect on the composite rate of complications (P=0.67) or on the percentage of joint-replacement procedures performed in high-risk patients (P=0.81).

CONCLUSIONS

In the first 2 years of the CJR program, there was a modest reduction in spending per hip- or knee-replacement episode, without an increase in rates of complications. (Funded by the Commonwealth Fund and the National Institute on Aging of the National Institutes of Health.)

From the Department of Health Policy and Management, Harvard T.H. Chan School of Public Health (M.L.B., A.M.E., E.J.O.), the Department of Health Care Policy, Harvard Medical School (A.W., J.M.M., D.C.G., A.M.), and the Department of Medicine, Brigham and Women's Hospital (M.L.B., J.M.M., A.M.E., E.J.O.) — all in Boston; and the Department of Medicine, Washington University School of Medicine in St. Louis, St. Louis (K.E.J.M.). Address reprint requests to Dr. Barnett at the Department of Health Policy and Management, Harvard T.H. Chan School of Public Health, Kresge Bldg., Rm. 411, 677 Huntington Ave., Boston, MA 02115, or at mbarnett@hsph.harvard.edu.

Drs. Barnett and Wilcock contributed equally to this article.

This article was published on January 2, 2019, and updated on April 18, 2019, at NEJM.org.

N Engl J Med 2019;380:252-62.
DOI: 10.1056/NEJMsa1809010

In April 2016, Medicare initiated Comprehensive Care for Joint Replacement (CJR), a mandatory bundled-payment model for inpatient replacement of the hip or knee.[1] In the CJR program, hospitals are held accountable for spending for an episode of care, which includes the index hospitalization for the procedure plus all spending (with minor exceptions specified by the Centers for Medicare and Medicaid Services) in the 90 days after discharge.[2] In contrast to the voluntary nature of other alternative payment models, CJR randomly assigns metropolitan statistical areas to mandatory participation.[3,4] Hospitals in areas that are randomly assigned to the CJR program are subject to bundled payments for all episodes of hip or knee replacement.

Like other bundled-payment programs,[5,6] CJR was designed to provide financial incentives for hospitals to reduce spending without compromising quality across an entire episode of care during the index hospitalization and after discharge. During a CJR episode, fee-for-service payments are made as usual to all providers (e.g., outpatient physicians or skilled nursing facilities). Participating hospitals then undergo an annual retrospective reconciliation process in which their average spending per episode is compared with a hospital-specific benchmark. Hospitals share savings with Medicare if spending falls below the benchmark or, starting in 2017, they pay a penalty if spending exceeds the target.[1,2] As with the accountable care organization programs in Medicare,[7,8] the savings or losses of hospitals are adjusted according to their performance in a mix of hip- or knee-replacement quality measures such as rates of complications.

Voluntary bundled-payment programs have been associated with either unchanged or reduced spending without deterioration in quality.[9-11] However, changes observed in these programs could be due to the selective participation of highly motivated hospitals and providers.[12,13] As compared with previous voluntary programs, CJR is an important advance because it features both a randomized design and mandatory participation. Evaluations of the first year of CJR showed modest although not always significant decreases in total spending, without changes in quality.[14,15] As CJR matures, it is unclear whether these savings will become larger and whether negative unintended consequences, such as hospitals declining to treat sicker patients whose care could potentially be more costly,[12,13] will become evident.

We compared changes in spending and quality in the first 2 years of the CJR program between areas that were randomly assigned to the new bundled-payment model ("treatment" areas) and control areas. We also evaluated any potentially unintended consequences such as the selection of healthier patients for hip or knee replacement or an increased volume of these procedures.

Methods

Study Population

We analyzed Medicare claims and enrollment data from 2015 through 2017 on Medicare fee-for-service beneficiaries who underwent inpatient primary hip- or knee-replacement procedures (diagnosis-related group [DRG] 469 or DRG 470 at discharge) at hospitals in one of the 196 metropolitan statistical areas that were eligible for participation in the CJR program (Fig. S1 in the Supplementary Appendix, available with the full text of this article at NEJM.org). Revisions of previous hip- or knee-replacement procedures were not included in the CJR program. The unit of analysis was an episode of hip or knee replacement, defined as the period between the date of admission for hospitalization and 90 days after discharge.

The CJR program began in April 2016. In our study, the 12-month period before the initiation of the CJR model included procedures performed between January 1 and December 31, 2015, and the 15-month period after the initiation of the CJR model included procedures that occurred between July 1, 2016, and September 30, 2017. We used data through December 31, 2017, to cover the 90-day period after hospital discharge for all hip- or knee-replacement procedures; this encompassed the period used by CJR to assess performance in years 1 and 2 of the program. In the primary analysis, we excluded episodes that began during a 6-month transition period from January 1 through June 30, 2016, since during this time hip- or knee-replacement episodes overlapped into the period after the initiation of the CJR model or occurred early after initiation of the program, when hospitals may have been adjusting to the new payment model.

For each patient who underwent a hip- or

knee-replacement procedure, including those who did and those who did not have a hip fracture, we included the first admission for joint replacement during the 2015–2017 period. We excluded episodes for patients with active end-stage renal disease and episodes with any additional hip- or knee-replacement procedures (e.g., sequential knee replacements) in the 90 days after discharge from the index hospitalization. We also excluded joint-replacement procedures performed in hospitals that participated in the Bundled Payments for Care Improvement (BPCI) program for hip- or knee-replacement procedures[16] and any episode of hip or knee replacement in which the patient was discharged to a skilled nursing facility or home health agency that participated in the BPCI program for hip- or knee-replacement procedures.

We applied one additional exclusion criterion that was not part of the CJR program. To assess coexisting conditions in the patients and to ensure that we captured all care within the episode, we limited our analyses to patients who were continuously enrolled in fee-for-service Medicare Part A and Part B for 12 months before the hip- or knee-replacement episode through 90 days after discharge or until death. A comparison of the overall Medicare population and the continuously enrolled population is provided in Table S1 in the Supplementary Appendix.

This study was approved by the institutional review board at Harvard Medical School. Informed consent was waived because the data were deidentified.

STUDY VARIABLES

Randomization

Of 380 metropolitan statistical areas in the United States, 196 were identified as being eligible for the CJR program on the basis of volume of hip- or knee-replacement procedures and participation in the BPCI program (Methods section A in the Supplementary Appendix). Medicare categorized the 196 eligible areas into strata according to the median population (two strata). Each population stratum was then differentiated according to quartile of spending per hip- or knee-replacement episode before the initiation of the CJR model (four strata), for a total of eight strata. Within each stratum, Medicare then randomly assigned areas to treatment or control groups. The probability that an area would be assigned to mandatory participation in the CJR program was higher among areas with higher spending before the initiation of the program than among those with lower spending. To address potential bias from regression to the mean that would otherwise result from the oversampling within higher-spending strata, areas were weighted to equalize the probability of assignment to the treatment group within each of the eight strata (Methods section A in the Supplementary Appendix).

The initial randomization was announced by Medicare in July 2015,[2] but in November 2015, Medicare cut the number of metropolitan statistical areas that were randomly assigned to the CJR intervention from 75 to 67, largely because of updates in hospital participation in the BPCI program (Methods section A in the Supplementary Appendix). Because these cuts were not random, we conducted an intention-to-treat analysis based on the initial randomization, in which all patients who underwent hip or knee replacement in hospitals in the initially randomly assigned 75 metropolitan statistical areas made up the treatment population and patients in the remaining 121 metropolitan statistical areas made up the control population.

Patient, Hospital, and Procedure Characteristics

From Medicare enrollment files, we determined the patients' age, sex, and race or ethnic group; the rural or urban categorization of their ZIP Code of residence[17]; the original reason for Medicare enrollment (i.e., disability or age); their Medicaid enrollment; and the presence of 27 chronic conditions from the Chronic Conditions Data Warehouse.[18] For each joint-replacement episode, we assessed the type of procedure (hip or knee replacement and total or partial replacement), the DRG, and whether the patient had a hip fracture.

Primary Outcomes

Our first primary outcome was institutional spending, which included Medicare payments to a hospital (inpatient or outpatient), post–acute care facility, or hospice, as well as spending on durable medical equipment (further details are provided in Methods section B in the Supplementary Appendix). We chose institutional spending as a primary outcome because it makes up

approximately 85% of all spending on hip- or knee-replacement episodes, it is the component of spending in which previous studies of bundled payments for hip or knee replacement have shown savings,[9,19] and data on noninstitutional spending (payments for physicians and other providers, ambulance services, and independent laboratories) were available for only a 20% random sample of Medicare beneficiaries (additional details are provided in Table S2 in the Supplementary Appendix).

Our second primary outcome was a composite measure of complications of hip or knee replacement that was developed and used by Medicare for public reporting.[20] In this measure, a complication was defined as any of several medical complications (e.g., pulmonary embolism) or surgical complications (e.g., joint infection) during the procedure or within 90 days after the admission date (full details are provided in Methods section B in the Supplementary Appendix). We included all patients in this quality measure.

Our third primary outcome was the percentage of patients undergoing hip- or knee-replacement procedures for whom there was at elevated risk of high overall spending in the joint-replacement episode. In the CJR model, hospitals have a financial incentive to selectively decline to treat sicker patients whose care may be more expensive, because the CJR program does not adjust the benchmark of each hospital for patient factors other than the presence of a hip fracture.[21] We estimated each patient's burden of illness for each episode by estimating a "risk score" that was based on predicted total spending per episode. Risk scores were estimated with the use of a linear regression model predicting total spending for each episode fitted with data incorporating patient characteristics from 2013 through 2014 (before the initiation of the CJR model) (Methods section C in the Supplementary Appendix). Using this model, we assigned each episode to a quartile-of-risk score and assessed the percentage of hip- or knee-replacement procedures performed in patients in the highest quartile.

Secondary Outcomes

We examined several secondary outcomes that are detailed in Methods section B in the Supplementary Appendix. Using claims from a 20% random sample of beneficiaries, we measured total Medicare spending per hip- or knee-replacement episode. Utilization measures included the use of post–acute care services according to facility type, length of stay in post–acute care facilities, and readmission or visits to an emergency department within 90 days after discharge. We also measured mortality at 90 days and complications excluding hip fractures, as used by Medicare for public reporting. We examined differential changes (i.e., the between-group differences in the change from the period before the CJR model) in the DRG used for hip- or knee-replacement episodes (469 or 470), the mean risk score, sociodemographic characteristics, and the presence of chronic coexisting conditions. To assess the contribution of differential changes in observable patient characteristics to our estimates, we also evaluated the results of our primary analysis with and without adjustment for characteristics of the patients and episodes.

STATISTICAL ANALYSIS

Our primary analysis involved a difference-in-differences approach. We fit a linear regression model at the hip- or knee-replacement episode level with adjustment for characteristics of the patients and procedures as well as for hospital and metropolitan statistical area random effects to control for time-invariant differences between treatment and control hospitals and areas, as well as fixed effects for each quarter of our study period (see Methods section D in the Supplementary Appendix). The key variable in the model was an interaction between the period after the initiation of the CJR model and an indicator for the procedure being performed in a treatment area; this describes the mean differential change in the outcome for episodes in treatment areas relative to those in control areas (i.e., the estimated effect of the CJR program). In a prespecified alternative modeling approach, we used hospital fixed effects with robust variance estimators at the metropolitan statistical area level (Methods section D in the Supplementary Appendix). In sensitivity analyses, we examined generalized linear models with a log link and mean proportional to variance function for continuous outcomes and logistic-regression models for binary outcomes. Finally, in post hoc analyses, we also compared outcomes in three 6-month periods from April through September 2016, October

Table 1. Characteristics of the Patients in 2015, before Implementation of the CJR Model.*

Variable	Treatment Group (N=102,089)	Control Group (N=143,824)
Mean age (yr)	74.5	74.3
Male sex (%)	35.9	36.0
DRG (%)		
DRG 469, hip or knee replacement with major complication or coexisting condition	5.6	5.1
DRG 470, hip or knee replacement without major complication or coexisting condition	94.4	94.9
Fracture (%)	16.2	15.0
Procedure (%)		
Total knee replacement	54.2	56.2
Total hip replacement	31.4	30.9
Race or ethnic group (%)†		
Non-Hispanic white	89.8	90.8
Non-Hispanic black	5.9	5.6
Hispanic	1.1	0.7
Asian	1.0	0.6
Other	2.2	2.2
Original reason for Medicare enrollment (%)		
Age >65 yr	84.2	84.1
Disability	15.7	15.8
End-stage renal disease‡	0.1	0.1
Eligible for Medicaid (%)	11.4	10.3
Urban residence (%)§	85.0	82.4
Previous inpatient care (%)	20.4	19.7
Previous post–acute care services (%)	7.7	7.2
Chronic conditions (mean no.)¶	7.1	7.0

* Weighted numbers of patients are shown. CJR denotes Comprehensive Care for Joint Replacement, and DRG diagnosis-related group.
† Race or ethnic group was determined from the Medicare Master Beneficiary Summary Files. Percentages may not total 100 because of rounding.
‡ Patients with end-stage renal disease were excluded from the payment program. However, some patients initially qualified for Medicare because of end-stage renal disease but were no longer classified as having this disease at the time of the hip or knee replacement.
§ Urban residence, which was determined with the use of the Health Resources and Services Administration Rural–Urban Commuting Area Codes database (www.depts.washington.edu/uwruca/index.php), was defined as residence in a "metropolitan" ZIP Code. Data were missing for 0.16% of episodes that primarily occurred in Puerto Rico.
¶ We assessed patients for the presence of 27 conditions from the Chronic Conditions Data Warehouse, which uses claims since 1999 to describe the accumulated chronic disease burden in Medicare beneficiaries (a list of conditions is provided in Methods section C in the Supplementary Appendix).

2016 through March 2017, and April through September 2017 (Methods section D in the Supplementary Appendix). In addition, we used publicly released data[1] on the net reconciliation payments made by Medicare to hospitals in the first 2 years of CJR to estimate the net savings or loss to Medicare (Methods section E in the Supplementary Appendix).

There were no significant differences in the trends between the treatment and control areas before the initiation of the CJR model for each outcome in 2015, the "pre-period" used in our models (Table S3 in the Supplementary Appendix). Because 2015 is a single year, we also compared trends before the initiation of the CJR model over a longer period (2011 through 2015) in a post hoc analysis (more details are provided in Methods section F in the Supplementary Appendix).

To adjust for multiple testing, we set the significance level for each of the three primary outcomes at less than 0.0167 (0.05÷3). We provide 95% confidence intervals, without P values, for exploratory estimates of secondary outcomes, which were not adjusted for multiple testing.

Our analytic protocol was prespecified and is available at NEJM.org. The protocol was published before we performed the analyses of data from the period after the initiation of the CJR model.[22]

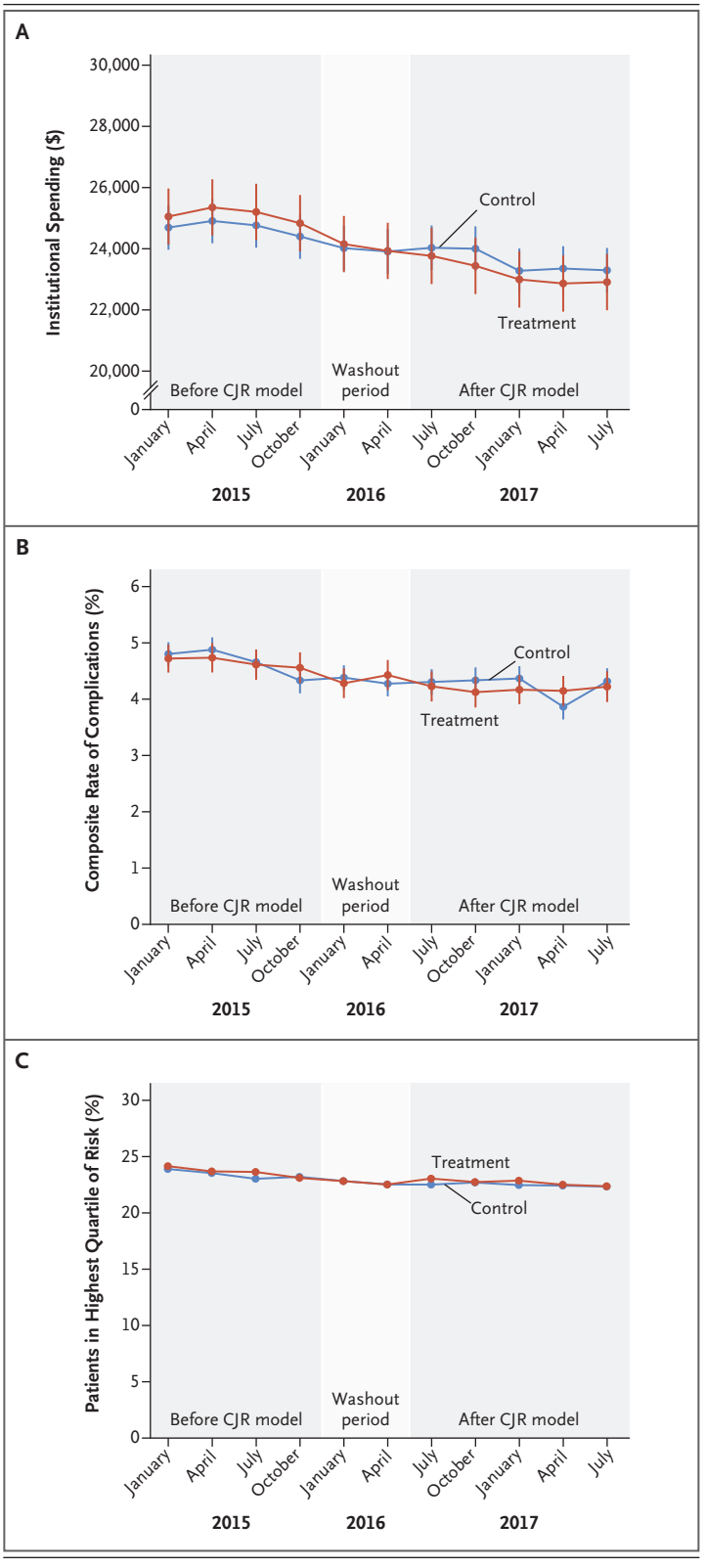
RESULTS

STUDY POPULATION

From 2015 through 2017, a total of 280,161 hip- or knee-replacement procedures were performed in 803 hospitals in treatment areas and 377,278 procedures were performed in 962 hospitals in control areas (unweighted). In the treatment areas, 7% of the procedures were performed in the 8 areas that were excluded after randomization. In 2015, before the initiation of the CJR model, there were differences between patient characteristics in the treatment versus control areas; for example, 11.4% of the episodes were for Medicaid-eligible patients in the treatment areas, as compared with 10.3% in the control areas (Table 1). There were meaningful differences in the characteristics of hospitals and counties in the treatment areas versus control areas in 2015 (Tables S4 and S5 in the Supplementary Appendix).

Figure 1. Adjusted Trends in Primary Outcomes, 2015–2017.

Shown are adjusted estimates for each of the three primary outcomes: institutional spending (Panel A), the rate of complications of hip or knee replacement (Panel B), and the percentage of patients undergoing these procedures who were in the top risk-score quartile (Panel C). All estimates are adjusted for hospital and metropolitan statistical areas as random effects. Estimates of institutional spending and complication rates are also adjusted for characteristics of the patients and episodes, as described in Methods section D in the Supplementary Appendix. The percentage of patients in the highest quartile of risk was not adjusted for characteristics of the patients and episodes because these characteristics are used to generate the patient risk score, which uses coefficients estimated from 2013–2014 data. In Panel A, the inset shows the same data on an enlarged y axis.



SPENDING AND UTILIZATION

Institutional spending on hip- or knee-replacement episodes decreased from $25,903 to $23,915 in the treatment group and from $24,596 to $23,238 in the control group (adjusted differential change between the treatment group and the control group, −$812; P<0.001), or a 3.1% differential decrease relative to mean spending in treatment areas before the CJR intervention (Fig. 1 and Table 2). This decrease in institutional spending grew over the 18-month period of CJR implementation from a differential change of −$541 in April through September 2016 (95% confidence interval [CI], −754 to −328) to −$860 in April through September 2017 (95% CI, −1,075 to −645) (Table S6 in the Supplementary Appendix).

The reduction in differential spending was driven largely by reduced spending at post–acute care facilities: skilled nursing facilities (adjusted differential change, −$527; 95% CI, −611 to −443) and inpatient rehabilitation facilities (adjusted differential change, −$227; 95% CI, −274 to −180). In the 20% sample of beneficiaries, total spending for hip- or knee-replacement procedures (including all professional fees in addition to payments to institutions) differentially decreased by $1,084 (95% CI, −1,409 to −760), or a 3.6% differential reduction (Table 3). On the basis of Medicare reconciliation payments of $872 per hip- or knee-replacement episode to hospitals

Table 2. Differential Changes in Primary Outcomes before and after Implementation of the CJR Model.*

Primary Outcome	Treatment Group			Control Group			Adjusted Difference-in-Differences Estimate (95% CI)†	P Value
	Average Episode before CJR	Average Episode after CJR	Unadjusted Difference	Average Episode before CJR	Average Episode after CJR	Unadjusted Difference		
Institutional spending (U.S. dollars)‡	25,903	23,915	−1,988	24,596	23,238	−1,358	−812 (−981 to −644)	<0.001
Composite rate of complications (%)§	4.72	4.15	−0.57	4.56	4.00	−0.56	−0.04 (−0.2 to 0.2)	0.67
Episodes with top quartile of patient risk (%)	25.30	23.60	−1.70	23.12	21.62	−1.50	−0.1 (−0.5 to 0.4)	0.81

* All estimates were adjusted for hospital and metropolitan statistical area random effects. Estimates of institutional spending and the rate of complications were also adjusted for characteristics of the patients and episodes, as described in Methods section D in the Supplementary Appendix. The outcome of the top quartile of patient risk was not adjusted for characteristics of the patients and episodes because these characteristics were used to generate the patient risk score, which uses coefficients estimated from 2013–2014 data. CI denotes confidence interval.

† Values for the difference between percentages are percentage points.

‡ A detailed definition of institutional spending is provided in Table S2 in the Supplementary Appendix.

§ According to the Medicare approach, each episode was classified as having no complications or one or more complications within 90 days after discharge following the index hospitalization.

from 2016 through 2017, the average net reduction in total Medicare spending per episode was $212, or a 0.7% relative decrease (Methods section E in the Supplementary Appendix). There was no significant differential change in the per capita volume of hip- or knee-replacement episodes between the treatment areas and control areas after the CJR intervention (Table S7 in the Supplementary Appendix).

The CJR program was associated with a 2.5-percentage-point (95% CI, −3.0 to −2.1) differential decrease (or a 5.9% relative decrease) in the percentage of patients discharged to post–acute care facilities and a differential reduction of 1.7 days in the length of stay in post–acute care facilities among those in any post–acute care facility (95% CI, −2.0 to −1.4) (Table 3). Estimates of these outcomes were not substantively different with the use of models with fixed effects, generalized linear models, or logistic-regression models (Table S8 in the Supplementary Appendix).

QUALITY OUTCOMES

The CJR program did not have a significant differential effect on the composite rate of complications (adjusted differential change in the percentage of episodes associated with a complication, −0.04%; P=0.67) (Fig. 1 and Table 2). In sensitivity analyses excluding admissions for hip fracture, results were similar (adjusted differential change in the percentage of episodes associated with a complication, −0.05%; 95% CI, −0.2 to 0.1) (Table 3). The program had no significant negative effect on the use of hospitals after discharge (inpatient, emergency department, or observation stay) or on mortality (Table 3).

RISK SELECTION

There was no significant differential change in our primary outcome for patient selection, in the change in the percentage of patients in the top risk-score quartile (differential change) (Fig. 1 and Table 2), or in the average patient risk score (Table S9 in the Supplementary Appendix). In treatment areas, there was a differential decrease in the percentage of patients who originally enrolled in Medicare because of disability (−0.6 percentage points; 95% CI, −1.0 to −0.2; or relative change of −3.8%) (Table S10 in the Supplementary Appendix). However, the estimated effect of the CJR program on spend-

ing per episode was reduced by only 2.5% after adjustments for observed characteristics of the patients before accounting for characteristics of the episodes (Table S11 in the Supplementary Appendix).

DISCUSSION

We found that through the first 2 years of the CJR program, the program modestly reduced payments per hip- or knee-replacement episode without any significant change in rates of complications. This decrease in payments grew over an 18-month period, which raises the possibility that CJR could lead to greater reductions in payments as hospitals adapt to the new payment model. The 3% reduction in payments was significantly offset by bonuses paid by Medicare to hospitals with spending below their CJR benchmark, although even after these payments, there was a small net savings.[1,23]

Since the CJR program is one of the only payment models in Medicare implemented as a mandatory randomized trial, it is an unusual experiment in payment reform. The mandatory participation in the CJR program generated considerable controversy, culminating in the Trump administration transitioning the program to a partly voluntary model as of March 2018.[24] Although the future of mandatory payment models is uncertain, the CJR program helps address the question of whether savings seen in previous evaluations of bundled-payment programs were attributable to the select nature of the hospitals that volunteered. Our findings suggest that the changes observed in voluntary programs may be echoed in mandatory programs.

Decreased Medicare spending on hip- and knee-replacement episodes at hospitals in the CJR program was nearly exclusively related to reductions in the use of post–acute care services in skilled nursing facilities and inpatient rehabilitation facilities. This is not surprising, because post–acute care services are a large and highly variable fraction of spending in hip- or knee-replacement episodes[25,26] and hospitals have strong financial incentives to reduce the frequency of post–acute care services. We did not find a negative effect on the rate of complications, readmissions, or death under the CJR program; therefore, it appears that hospitals may have successfully identified patients who are at the margin of needing post–acute care services who could instead be safely discharged home with home health services. However, our measures of quality do not include important patient-centered measures such as functional status, pain, and overall satisfaction. There may have been a negative effect on these dimensions of quality that we did not observe.

Our results are consistent with previous data showing that savings in bundled-payment models[9,14,15,19,25,27] and other alternative payment models[28,29] have been concentrated in changing the use of post–acute care services. Post–acute care services may be the easiest target for hospitals to decrease episode-level spending because it is often unclear when these services are beneficial or what intensity of post–acute care is most appropriate.[30-32]

One concern about current bundled-payment programs is that they create a financial incentive to treat healthier patients rather than those who are sicker and whose care may be more costly. There has been inconsistent evidence on risk selection in previous evaluations of voluntary bundling and the CJR program.[9,33] Although we did not see any substantive changes in our outcome of primary risk selection, in treatment areas, we found evidence of differential reductions in the percentage of disabled patients undergoing hip- or knee-replacement procedures. Adjustment for these and other observable characteristics of the patients had a minor effect on our estimates of savings. However, we could not examine whether changes in other unobserved risk factors for high spending after surgery may have contributed to our results. Close monitoring for risk selection under the CJR program is warranted.

Our study has several limitations. Our conclusions regarding the effects of bundled payments may not be generalizable beyond hip- or knee-replacement procedures,[34] and our evaluation covers only the first 2 years of the program. However, because the CJR model transitioned to a partly voluntary model 3 months after the end of our study period, our evaluation encompasses all but 2 months when the model was mandatory. Our evaluation focused on payments from Medicare and did not assess the overall financial effect on hospitals. Finally, our estimate of savings may be an underestimate of the true effect of bundled payments because we included hip- or

Table 3. Differential Changes in Spending and Quality Outcomes before and after Implementation of the CJR Model.*

Outcome	Treatment Group			Control Group			Adjusted Difference-in-Differences Estimate (95% CI)†
	Average Episode before CJR	Average Episode after CJR	Unadjusted Difference	Average Episode before CJR	Average Episode after CJR	Unadjusted Difference	
Spending (U.S. dollars)							
Total spending in 20% of sample‡	30,504	27,950	−2,554	28,836	27,193	−1,643	−1,084 (−1,409 to −760)
For index hospitalization	14,733	14,605	−128	14,326	14,086	−240	−67 (−179 to 44)
For other inpatient hospitalization (e.g., readmissions)	1,514	1,481	−33	1,395	1,390	−5	−22 (−88 to 45)
For skilled nursing facility	5,252	3,898	−1,354	4,587	3,787	−800	−527 (−611 to −443)
For inpatient rehabilitation facility	1,261	877	−384	1,201	1,030	−171	−227 (−274 to −180)
For long-term care hospital	86	56	−30	113	76	−37	9 (−13 to 30)
For home health agency	2,056	1,988	−68	1,963	1,857	−106	−4 (−22 to 14)
For hospice	68	72	4	63	69	6	0 (−8 to 8)
For outpatient facility§	792	831	39	811	833	22	22 (2 to 42)
For professional services in 20% of sample	4,182	4,125	−57	4,017	3,975	−42	−37 (−93 to 18)
For durable medical equipment	141	107	−34	138	111	−27	−5 (−10 to 1)
Utilization of hospital and post–acute care services							
Mean length of stay, index hospitalization (days)	4.15	3.81	−0.34	4.04	3.72	−0.3	−0.02 (−0.04 to −0.004)
Discharge to post–acute care facility (%)¶	42.2	31.6	−10.6	41.3	33.1	−8.2	−2.5 (−3.0 to −2.1)
Discharge to home health agency (%)	34.6	38.5	3.9	33.0	33.9	0.9	2.5 (2.0 to 2.9)
Mean length of stay, post–acute care facility (days)‖	22.65	20.62	−2.03	21.13	20.67	−0.5	−1.7 (−2.0 to −1.4)
Mean length of stay, post–acute care facility, not including inpatient rehabilitation facility (days)‖	23.44	20.79	−2.65	21.71	21.06	−0.6	−2.1 (−2.4 to −1.8)
Home health agency (mean no. of episodes)	12.74	11.55	−1.19	12.76	12.47	−0.3	−0.9 (−1.0 to −0.8)
Composite of complications and mortality (%)							
Medicare-defined complications, not including fractures**	2.4	2.2	−0.2	2.5	2.2	−0.3	−0.05 (−0.2 to 0.1)
90-Day mortality	2.6	2.2	−0.4	2.3	2.1	−0.2	−0.06 (−0.2 to 0.1)
Hospital use within 90 days after discharge (%)							
Inpatient readmission for any cause	9.7	9.1	−0.6	9.2	9.5	0.3	−0.6 (−0.9 to −0.2)
ED visit without admission	13.5	13.7	0.2	13.8	13.9	0.1	0.3 (−0.1 to 0.7)
Observation stay without admission	2.1	2.2	0.1	2.3	2.4	0.1	0.05 (−0.1 to 0.2)
Any ED, observation, or inpatient visit	22.1	21.6	−0.5	21.8	22.2	0.4	−0.3 (−0.8 to 0.1)

* Values for the difference between percentages are percentage points. All estimates were adjusted for hospital and metropolitan statistical area random effects as well as for patient and episode characteristics, as described in Methods section D of the Supplementary Appendix. ED denotes emergency department.

† The 95% confidence intervals are not adjusted for multiple testing and should be interpreted with caution, since the analyses are exploratory.

‡ Total spending includes both institutional and noninstitutional spending (Table S2 in the Supplementary Appendix), or both Medicare Part A and Part B spending. This spending estimate is based on a 20% random sample of hip- or knee-replacement episodes because data on spending for professional services were available only for 20% of the sample of Medicare beneficiaries.

§ Spending on outpatient facilities includes payments to hospitals for visits to physician offices and for imaging or laboratory tests provided in facilities attached to or affiliated with a hospital. It does not include separate payments to physicians for their services.

¶ Post–acute care facilities include all inpatient post–acute care facilities (largely skilled nursing facilities) as well as inpatient rehabilitation facilities and long-term acute care hospitals.

‖ The length of stay was estimated only among patients who used post–acute care services (with or without use of an inpatient rehabilitation facility), and the number of episodes was estimated only among patients who used a home health agency. Length of stay for post–acute care services was estimated with or without inclusion of inpatient rehabilitation facilities because these facilities are not paid according to the length of stay.

** The composite rate of complications (no complications or one or more complications within 90 days after surgery) excludes hip- or knee-replacement episodes due to hip fracture. This is the approach used by Medicare to estimate this measure for public quality reporting.

knee-replacement episodes performed in eight metropolitan statistical areas (accounting for 7% of treatment episodes) that were originally randomly assigned to participate in the CJR program but then were subsequently excluded.

In conclusion, we found that in the first 2 years of the CJR program, mandatory bundled-payment models for joint replacement led to modest decreases in spending per episode. These decreases were driven by a reduced use of post–acute care services, without any major change in the rate of complications.

A data sharing statement provided by the authors is available with the full text of this article at NEJM.org.

Supported by grants from the Commonwealth Fund (to Drs. Wilcock, Epstein, Joynt Maddox, and Mehrotra) and the National Institute on Aging of the National Institutes of Health (K23 AG058806, to Dr. Barnett; and P01 AG032952, to Dr. McWilliams).

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

REFERENCES

1. Center for Medicare & Medicaid Innovation. Comprehensive Care for Joint Replacement model (https://innovation.cms.gov/initiatives/CJR).
2. Medicare program: Comprehensive Care for Joint Replacement payment model for acute care hospitals furnishing lower extremity joint replacement services. Fed Regist 2015;80(134):41198-316 (https://www.federalregister.gov/documents/2015/07/14/2015-17190/medicare-program-comprehensive-care-for-joint-replacement-payment-model-for-acute-care-hospitals).
3. Mechanic RE. Mandatory Medicare bundled payment — is it ready for prime time? N Engl J Med 2015;373:1291-3.
4. Wadhera RK, Yeh RW, Joynt Maddox KE. The rise and fall of mandatory cardiac bundled payments. JAMA 2018;319:335-6.
5. Porter ME, Lee TH. The strategy that will fix health care. Harvard Business Review. October 2013 (https://hbr.org/2013/10/the-strategy-that-will-fix-health-care).
6. Press MJ, Rajkumar R, Conway PH. Medicare's new bundled payments: design, strategy, and evolution. JAMA 2016;315:131-2.
7. Center for Medicare & Medicaid Innovation. Next generation ACO model (https://innovation.cms.gov/initiatives/Next-Generation-ACO-Model/).
8. Centers for Medicare & Medicaid Services (CMS). Medicare program: Medicare Shared Savings Program: accountable care organizations: final rule. Fed Regist 2015;80(110):32691-845.
9. Dummit LA, Kahvecioglu D, Marrufo G, et al. Association between hospital participation in a Medicare bundled payment initiative and payments and quality outcomes for lower extremity joint replacement episodes. JAMA 2016;316:1267-78.
10. 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;53:632-48.
11. Center for Medicare and Medicaid Services. Medicare Acute Care Episode (ACE) demonstration (https://innovation.cms.gov/initiatives/ACE/).
12. Fisher ES. Medicare's Bundled Payment Program for joint replacement: promise and peril? JAMA 2016;316:1262-4.
13. Pronovost PJ, Miller J, Newman-Toker DE, Ishii L, Wu AW. We should measure what matters in bundled payment programs. Ann Intern Med 2018;168:735-6.
14. 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;320:892-900.
15. The Lewin Group. CMS Comprehensive Care for Joint Replacement Model: performance year 1 evaluation report. 2018 (https://innovation.cms.gov/Files/reports/cjr-firstannrpt.pdf).
16. Centers for Medicare & Medicaid Services. Bundled Payments for Care Improvement (BPCI) initiative: general information (https://innovation.cms.gov/initiatives/bundled-payments/).
17. Morrill R, Cromartie J, Hart G. Rural-urban commuting code database (http://depts.washington.edu/uwruca/index.php).
18. Chronic Conditions Data Warehouse. 2014 (https://www.ccwdata.org/).
19. Navathe AS, Troxel AB, Liao JM, et al. Cost of joint replacement using bundled payment models. JAMA Intern Med 2017;177:214-22.
20. Medicare.gov. Complication rate for hip/knee replacement patients (https://www.medicare.gov/hospitalcompare/Data/Surgical-Complications-Hip-Knee.html).
21. HealthPayerIntelligence. CMS bundled payment models lead to greater patient selectivity. 2017 (https://healthpayerintelligence.com/news/cms-bundled-payment-models-lead-to-greater-patient-selectivity).
22. Barnett ML. Comprehensive Care for Joint Replacement Program evaluation. Open Science Framework. 2018 (https://osf.io/2pdtf/).
23. Navathe AS, Liao JM, Shah Y, et al. Characteristics of hospitals earning savings in the first year of mandatory bundled payment for hip and knee surgery. JAMA 2018;319:930-2.
24. CMS finalizes changes to the Comprehensive Care for Joint Replacement Model, cancels Episode Payment Models and Cardiac Rehabilitation Incentive Payment Model. Press release of the Centers for Medicare & Medicaid Services, November 30, 2017 (https://www.cms.gov/Newsroom/MediaReleaseDatabase/Press-releases/2017-Press-releases-items/2017-11-30.html).
25. Tsai TC, Joynt KE, Wild RC, Orav EJ, Jha AK. Medicare's bundled payment initiative: most hospitals are focused on a few high-volume conditions. Health Aff (Millwood) 2015;34:371-80.
26. Huckfeldt PJ, Mehrotra A, Hussey PS. The relative importance of post-acute care and readmissions for post-discharge spending. Health Serv Res 2016;51:1919-38.
27. Jubelt LE, Goldfeld KS, Chung WY, Blecker SB, Horwitz LI. Changes in discharge location and readmission rates under Medicare bundled payment. JAMA Intern Med 2016;176:115-7.
28. McWilliams JM, Gilstrap LG, Stevenson DG, Chernew ME, Huskamp HA, Grabowski DC. Changes in postacute care in the Medicare Shared Savings Program. JAMA Intern Med 2017;177:518-26.
29. McWilliams JM, Hatfield LA, Chernew ME, Landon BE, Schwartz AL. Early performance of accountable care organizations in Medicare. N Engl J Med 2016;374:2357-66.
30. Kane RL. Finding the right level of posthospital care: "We didn't realize there was any other option for him." JAMA 2011;305:284-93.
31. Hoverman C, Shugarman LR, Saliba D, Buntin MB. Use of postacute care by nursing home residents hospitalized for stroke or hip fracture: how prevalent and to what end? J Am Geriatr Soc 2008;56:1490-6.
32. Buntin MB, Colla CH, Deb P, Sood N, Escarce JJ. Medicare spending and outcomes after postacute care for stroke and hip fracture. Med Care 2010;48:776-84.

33. Navathe AS, Liao JM, Dykstra SE, 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;320:901-10.
34. Joynt Maddox KE, Orav EJ, Zheng J, Epstein AM. Evaluation of Medicare's bundled payments initiative for medical conditions. N Engl J Med 2018;379:260-9.

Western Scrub Jay Taking Off Calvin Kunin, M.D.