

JAMA | Original Investigation

# 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

Amy Finkelstein, PhD; Yunan Ji, BA; Neale Mahoney, PhD; Jonathan Skinner, PhD

**IMPORTANCE** Bundled payments are an increasingly common alternative payment model for Medicare, yet there is limited evidence regarding their effectiveness.

**OBJECTIVE** To report interim outcomes from the first year of implementation of a bundled payment model for lower extremity joint replacement (LEJR).

**DESIGN, SETTING, AND PARTICIPANTS** As part of a 5-year, mandatory-participation randomized trial by the Centers for Medicare & Medicaid Services, eligible metropolitan statistical areas (MSAs) were randomized to the Comprehensive Care for Joint Replacement (CJR) bundled payment model for LEJR episodes or to a control group. In the first performance year, hospitals received bonus payments if Medicare spending for LEJR episodes was below the target price and hospitals met quality standards. This interim analysis reports first-year data on LEJR episodes starting April 1, 2016, with data collection through December 31, 2016.

**EXPOSURE** Randomization of MSAs into the CJR bundled payment model group (75 assigned; 67 included) or to the control group without the CJR model (121 assigned; 121 included). Instrumental variable analysis was used to evaluate the relationship between inclusion of MSAs in the CJR model and outcomes.

**MAIN OUTCOMES AND MEASURES** The primary outcome was share of LEJR admissions discharged to institutional postacute care. Secondary outcomes included the number of days in institutional postacute care, discharges to other locations, Medicare spending during the episode (overall and for institutional postacute care), net Medicare spending during the episode, LEJR patient volume and patient case mix, and quality-of-care measures.

**RESULTS** Among the 196 MSAs and 1633 hospitals, 131 285 eligible LEJR procedures were performed during the study period (mean volume, 110 LEJR episodes per hospital) among 130 343 patients (mean age, 72.5 [SD, 0.91] years; 65% women; 90% white). The mean percentage of LEJR admissions discharged to institutional postacute care was 33.7% (SD, 11.2%) in the control group and was 2.9 percentage points lower (95% CI, -4.95 to -0.90 percentage points) in the CJR group. Mean Medicare spending for institutional postacute care per LEJR episode was \$3871 (SD, \$1394) in the control group and was \$307 lower (95% CI, -\$587 to -\$27) in the CJR group. Mean overall Medicare spending per LEJR episode was \$22 872 (SD, \$3619) in the control group and was \$453 lower (95% CI, -\$909 to \$3) in the CJR group, a statistically nonsignificant difference. None of the other secondary outcomes differed significantly between groups.

**CONCLUSIONS AND RELEVANCE** In this interim analysis of the first year of the CJR bundled payment model for LEJR among Medicare beneficiaries, MSAs covered by CJR, compared with those that were not, had a significantly lower percentage of discharges to institutional postacute care but no significant difference in total Medicare spending per LEJR episode. Further evaluation is needed as the program is more fully implemented.

**TRIAL REGISTRATION** ClinicalTrials.gov Identifier: [NCT03407885](https://clinicaltrials.gov/ct2/show/study/NCT03407885); American Economic Association Registry Identifier: [AEARCTR-0002521](https://www.aearegistry.org/registries/0002521)

JAMA. 2018;320(9):892-900. doi:10.1001/jama.2018.12346

← Editorial page 877

← Related article page 901

+ Supplemental content

+ CME Quiz at [jamanetwork.com/learning](https://jamanetwork.com/learning)

**Author Affiliations:** Department of Economics and J-PAL North America, Massachusetts Institute of Technology, Cambridge (Finkelstein); National Bureau of Economic Research, Cambridge, Massachusetts (Finkelstein, Mahoney, Skinner); Graduate School of Arts and Sciences, Harvard University, Cambridge, Massachusetts (Ji); Booth School of Business, University of Chicago, Chicago, Illinois (Mahoney); Department of Economics, Dartmouth College, Hanover, New Hampshire (Skinner).

**Corresponding Author:** Amy Finkelstein, PhD, Department of Economics, Massachusetts Institute of Technology, 77 Massachusetts Ave, Bldg E52, Room 442, Cambridge, MA 02139 (afink@mit.edu).

The shift toward alternative payment models in Medicare is an important trend in US health care. By 2016, 30% of traditional Medicare reimbursement had been shifted from fee-for-service (FFS) models to alternative payment models.<sup>1</sup> Bundled payments are one of the leading alternative payment models. Under bundled payments, health care organizations (such as hospitals, physician groups, and postacute care providers) receive a single “bundled” payment for all services related to a specific treatment (eg, hip replacement). By holding multiple parties jointly accountable for quality and costs, bundled payments may encourage coordination of care and reduce unnecessary utilization. However, because these parties are paid a fixed amount irrespective of the volume of services, they may also respond by reducing necessary care or by trying to treat only healthier patients with lower expected costs.<sup>2</sup>

Bundled payments have been widely touted for their potential to have a substantial, positive effect on health care delivery,<sup>3-7</sup> yet there is limited rigorous evidence on their effects. Most studies of bundled payments have been observational, focusing on the experience of a small number of hospitals that voluntarily participated. These studies have tended to find large savings,<sup>8-12</sup> but voluntary participation makes separating treatment from selection effects difficult,<sup>13</sup> and the small number of participating hospitals raises concerns about generalizability.

To address these gaps in scientific knowledge, this study took advantage of the Centers for Medicare & Medicaid Services' (CMS's) random assignment of some metropolitan statistical areas (MSAs) to a Medicare bundled payment model for lower extremity joint replacement (LEJR)—ie, hip and knee replacement. In 2014, there were 486 249 LEJR procedures, accounting for \$6.2 billion of Medicare inpatient spending (4.52% of all inpatient Medicare spending).<sup>14,15</sup> CMS designed and randomly assigned the 5-year, nationwide bundled payment program for LEJR, known as Comprehensive Care for Joint Replacement (CJR), which began on April 1, 2016. Unlike previous bundled payment programs, participation was mandatory for all covered hospitals. The purpose of this study was to analyze results from an interim analysis of the first performance year of this 5-year program.

## Methods

The primary objective of this study was to evaluate interim outcomes during the first year of implementation of a bundled payment model for LEJR, focusing on discharge to institutional postacute care following joint replacement hospitalization episodes.

Institutional review board (IRB) exemption was obtained from the Massachusetts Institute of Technology's Committee on the Use of Humans as Experimental Subjects (COUHES 1710117275) and IRB approval for Medicare data analysis was obtained from Dartmouth's IRB (15475). CMS did not require informed consent for patients in CJR; both IRBs waived informed consent for the analysis of the Medicare claims data.

## Key Points

**Question** What was the change in discharge to institutional postacute care after lower extremity joint replacement episodes among Medicare beneficiaries following implementation of the Comprehensive Care for Joint Replacement (CJR) bundled payments in 2016?

**Findings** In this interim analysis of the first year of a 5-year randomized trial of 75 metropolitan statistical areas (MSAs) that were assigned the bundled payment model and 121 control MSAs that were not, the mean percentage of patient discharges to institutional postacute care was 33.7% in the control group and was 2.9 percentage points lower in MSAs covered by the CJR model, a significant difference.

**Meaning** These interim findings suggest that CJR may reduce institutional postacute care following lower extremity joint replacement episodes among Medicare beneficiaries, although further evaluation is needed as the program is fully implemented over time.

## Study Design

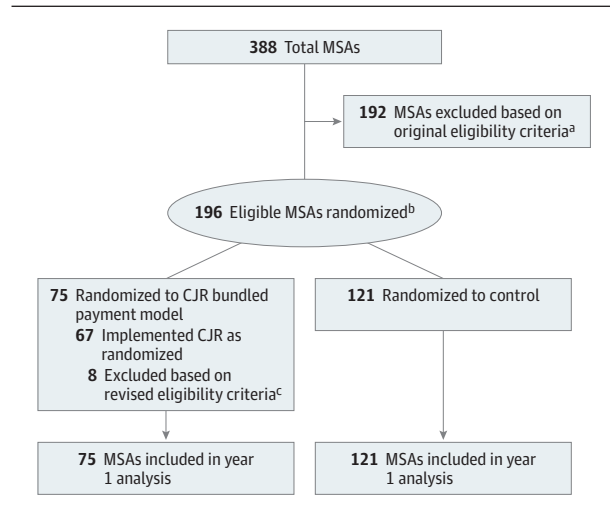
In July 2015, CMS publicly announced in the *Federal Register* its exclusion criteria and randomization procedure for selecting the 196 eligible MSAs<sup>16</sup>; MSAs were excluded primarily because of low LEJR discharge volume. Within eligible MSAs randomized to CJR, hospitals were required to participate in CJR if they were paid under prospective payment and not already participating in Model 1 or Phase 2 (Models 2 or 4) of the Bundled Payments for Care Improvement Initiative (BPCI), a preexisting Medicare voluntary bundled payments model for LEJR.

Within eligible hospitals, the patient inclusion criteria included Medicare Part A and Part B coverage, no readmission during the episode for LEJR, and no death during the episode. Applying the hospital and patient exclusion criteria to all Medicare FFS LEJR episodes in eligible MSAs in 2016, we estimated that 75% of episodes would be covered by CJR if the MSA was selected for treatment; section 1 of eAppendix 1 and eTable 1 in [Supplement 1](#) provide more detail on the eligibility criteria and this estimate.

The 196 eligible MSAs were divided into 8 strata based on quartile of historical LEJR payments and above- vs below-median MSA population. CMS set different treatment probabilities (ie, probability of selection for CJR) for MSAs by strata (ranging from 30% to 45%); MSAs with higher historical LEJR spending had higher treatment probabilities. CMS performed the randomization in SAS Enterprise Guide version 7.1 (SAS Institute Inc) using the PROC SURVEYSELECT statement with METHOD=SRS.

In July 2015, CMS publicly announced that based on this randomization, 75 MSAs were initially assigned to treatment and 121 to control ([Figure](#)). Following prior independent analyses of government-implemented randomization protocols,<sup>17,18</sup> we verified via simulation that we could reproduce the randomization procedure to within statistical sampling error (section 1 of eAppendix 1 and eTable 2 in [Supplement 1](#)).

**Figure. MSA Eligibility and Randomization in a 5-Year Randomized Trial of a Mandatory Medicare Bundled Payment Program for Lower Extremity Joint Replacement (LEJR) Episodes**



eTable 1 in the Supplement provides more details on the eligibility criteria and randomization process. MSA indicates metropolitan statistical area; CJR, Comprehensive Care for Joint Replacement program.

<sup>a</sup> Original eligibility criteria were (1) at least 400 LEJR episodes in the baseline period between July 1, 2013, and June 30, 2014; (2) at least 400 non-Bundled Payments for Care Improvement Initiative (BPCI) LEJRs in the baseline period; (3) at least 50% of LEJR episodes in the baseline period were non-BPCI; and (4) at least 50% of otherwise eligible LEJR episodes not in Maryland hospitals. For eligibility criterion 2, BPCI participation was defined as hospitals participating in BPCI model 1 and phase 2 of BPCI models 2 or 4 as of July 1, 2015. For eligibility criterion 3, BPCI participation was defined in 2 steps: first, less than 50% of potentially eligible LEJR episodes were in hospitals participating in phase 2 of BPCI models 2 or 4 as of July 1, 2015; second, less than 50% of LEJR referrals to skilled nursing facility or home health agency services were made up of skilled nursing facilities or home health agencies participating in BPCI model 3 as of July 1, 2015.

<sup>b</sup> The 196 eligible MSAs were divided into 8 strata based on the full interaction of (1) average wage-adjusted historical LEJR episode payment, grouped into quartiles, and (2) MSA population size, grouped into above and below median. Randomization occurred within strata. Treatment probabilities varied within the payment quartiles: 30% in the first quartile (lowest payment), 35% in the second, 40% in the third, and 45% in the fourth (highest payment).

<sup>c</sup> After randomization took place, the Centers for Medicare & Medicaid Services received comments that the original eligibility criteria did not take into account providers that entered into phase 2 of BPCI by October 1, 2015, which was the final quarter a phase 1 BPCI participant could transition into phase 2. CMS therefore revised the definition of BPCI participation in original eligibility criteria 2 and 3. The revised eligibility criterion 2 defined BPCI participation hospitals participants as of October 1, 2015, instead of as of July 1, 2015, and also included episodes associated with a physician who was in a physician group practice in phase 2 of BPCI model 2 as of October 1, 2015. Similarly, the revised eligibility criterion 3 defined BPCI participation based on the list of BPCI-participating hospitals, skilled nursing facilities, and home health agencies as of October 1, 2015, instead of as of July 1, 2015. The revised eligibility criteria resulted in exclusion of 8 MSAs from the CJR group, resulting in a final 67 MSAs in the CJR group. CMS did not announce which MSAs would have been excluded from the control group based on the revised criteria.

In November 2015, CMS publicly updated the MSA exclusion criteria in response to comments that the original criteria did not take into account hospitals and physician group practices that entered into Phase 2 BPCI by October 1, 2015; as a result, 8 MSAs were excluded from the treatment group without a corresponding set of MSAs excluded from the control group.<sup>19</sup>

## Intervention

The CJR is a Medicare bundled payment model for LEJR that holds acute care hospitals financially responsible for Medicare spending over the entire episode of care. An episode begins with a hospital stay with a discharge in 1 of 2 included diagnosis related groups (DRGs) (MS-DRGs 469 and 470) and ends 90 days after discharge. The CJR was introduced in April 2016 and designed to last for 5 years.

Under CJR, hospitals face financial incentives to reduce Medicare FFS spending and to maintain or increase quality. At the end of each year, hospitals that (1) have per-episode Medicare FFS spending below the target price (set by CMS based on historical hospital and regional episode spending and the reason for admission) and (2) have met a minimum quality standard (5 of 20 on a composite quality score) receive “shared savings” from CMS for the difference between the target price and spending up to a stop-gain amount (ie, maximum bonus payment to the hospital), with higher scores making them eligible for greater savings. Hospitals that have FFS spending of more than the target price are responsible for paying the difference up to a stop-loss amount (ie, maximum financial penalty to the hospital). The upside and downside risks increase over time. In the first year, the stop gain was +5% and there was no downside risk; by the fifth year, the stop gain and stop loss were each scheduled to be 20% of the target price.<sup>19,20</sup>

The use of random assignment by a government agency such as CMS is rare but not unprecedented and can be valuable for scientific research. For example, the state of Oregon used a random lottery to expand Medicaid coverage, enabling academic research on the Oregon Health Insurance Experiment.<sup>17,18,21-23</sup>

## Data and Outcome Measures

We studied the first performance year of CJR, which includes episodes that begin on or after April 1, 2016, and end no later than December 31, 2016. Specifically, we analyzed episodes starting between April 1, 2016, and September 15, 2016; the end date was chosen so that all episodes would fall within the performance year (given a mean length of stay for an LEJR admission of 3.1 days for DRG 470 and 7.0 days for DRG 469).<sup>24</sup>

We used Medicare FFS claims data for 100% of enrollees from 2012-2014 and 2016 in the 196 eligible MSAs. We limited the sample to episodes that would have been covered by CJR if the MSA were included in the treatment group. We omitted data from 2015 because treatment MSAs were announced midway through 2015 and behavior was potentially affected during that year. We also used Hospital Compare data<sup>25</sup> to construct an estimate of the targeted quality measure, and data on hospital-specific end-of-year reconciliation (ie, bonus) payments.<sup>20</sup> Section 2 of eAppendix 1 in Supplement 1 provides more detail on data and outcomes.

The primary outcome was the share of LEJR admissions discharged to institutional postacute care—these are skilled nursing facilities, long-term care hospitals, and inpatient rehabilitation facilities. Existing observational studies<sup>11,12</sup> suggested that this would be the primary margin of adjustment,

and this was a margin where power calculations suggested that reasonably sized effects could be detected (prespecified analysis plan available in eAppendix 2 in Supplement 1).

Secondary outcomes included the number of days in institutional postacute care during the episode, discharges to other locations, Medicare FFS spending during the episode (both overall and for institutional postacute care), and net Medicare spending during the episode, which adds to Medicare FFS spending any reconciliation payments made to a treatment hospital under CJR.

Other secondary outcomes included LEJR patient volume, patient comorbidity severity, and several quality measures. Patient comorbidity severity was measured by the Elixhauser Comorbidity Index, which is the sum of 31 different comorbidity indicators.<sup>26,27</sup> The targeted quality measure was a modified composite quality score, derived from measures of total hip/knee arthroplasty complication rates and patient experience ratings (see section 2 of eAppendix 1 in Supplement 1 for more detail). The score ranges from 0 to 18, with higher numbers indicating better quality. Because the score in the first year was based almost entirely on data from prior to the introduction of CJR,<sup>20</sup> we noted in the preanalysis plan that we did not expect it to be affected (eAppendix 2 in Supplement 1). We therefore also examined nontargeted quality measures: the 90-day emergency department visit rate, a quality measure used in prior analyses of voluntary bundled payments for LEJR<sup>11,12</sup>; the 90-day all-cause readmission rate, a standard quality measure previously used for LEJR<sup>11,28</sup>; and the 90-day complication rate for total hip and total knee arthroplasty, a part of the targeted quality measure that we could observe for admissions during the study period.

### Statistical Analyses

The analyses were conducted at the MSA level. Not all randomly assigned MSAs were included in CJR (Figure). As prespecified (eAppendix 2 in Supplement 1), we therefore used a standard instrumental variable approach<sup>29</sup> to evaluate the relationship between inclusion in CJR and the outcomes; in the first-stage regression, assignment to CJR was used as an instrument for inclusion in CJR, and in the second-stage regression, inclusion in CJR was related to outcomes. As prespecified, all regressions included strata fixed effects because treatment probabilities varied by strata, and controlled for 2 years of lags of the dependent variable (specifically, 2013 and 2014) to improve statistical power. Based on historical data, we estimated the analysis to have power to detect a 2-percentage-point reduction in the primary outcome (2-sided  $\alpha = .05$  at 80% power). All analyses were prespecified (prior to obtaining postintervention data) other than analysis of the Elixhauser Comorbidity Index, which was requested during the review process, and the specific breakdown of discharges to noninstitutional postacute care, which was done for ease of exposition. Because of the strata sampling design, comparison of raw means of the treatment and control MSAs was not appropriate; instead, values for the control MSAs and the estimated differences (using the instrumental variable approach) between bundled payment

and control MSAs are reported. Raw comparisons of means between original bundled payment and control MSAs—separately by strata—are presented in section 2 of eAppendix 1 in Supplement 1. Because we did not adjust for multiple testing of secondary outcomes, secondary outcome analyses should be considered exploratory.

A number of sensitivity and additional analyses were performed. Results were analyzed without controlling for the lags of the dependent variable, and intention-to-treat analyses were conducted to compare outcomes for the 75 MSAs originally assigned to treatment with outcomes for the 121 control MSAs. Additional detailed analyses of the number and type of patients were also reported. Section 2 of eAppendix 1 in Supplement 1 provides more detail.

All statistical analyses were conducted using Stata version 15.1 (StataCorp), with 2-tailed tests with a statistical significance threshold of  $P < .05$ .

## Results

### Study Population

Of the 196 eligible MSAs, 75 were initially assigned to treatment and 121 to control; 8 of the 75 were subsequently excluded from CJR so that 67 MSAs were covered by CJR. The eFigure in Supplement 1 shows the geographic distribution. Assignment to CJR thus increased the chance of being included in CJR by 89.1 percentage points (95% CI, 81.8-96.4 percentage points;  $P < .001$ ).

In 2016, 167 hospitals were excluded from treatment or control MSAs because of preexisting BPCI participation, leaving 1647 eligible hospitals; 1633 of these hospitals had an eligible episode. Table 1 describes the CJR-eligible study population at the MSA level. On average, among eligible MSAs, the patient study population was 90% white and 65% female, with a mean age of 72.5 (SD, 0.91) years and a mean Elixhauser Comorbidity Index of 2.4 (SD, 0.29). Eligible hospitals had a mean of 289.1 beds and performed a mean of 20.4 CJR-eligible LEJR procedures per month over the study period (April-September 2016); 25.3% were for-profit hospitals and 9.3% were teaching hospitals. Among eligible MSAs, the mean number of acute care hospitals was 14.8 and the mean number of institutional postacute care providers was 50.9.

### Balance

Prior to randomization, characteristics were balanced across control and treatment MSAs. Table 2 shows balance at baseline (2014) for the outcome variables; an *F* test failed to reject equality of all of the outcomes ( $P = .54$ ). eTable 3 in Supplement 1 shows balance on MSA demographics and eTables 4 and 5 in Supplement 1 show raw means for each stratum.

### Health Care Use and Spending

Table 3 shows results of the instrumental variable analysis of the relationship between inclusion of MSAs in the CJR model and health care use and Medicare spending. For the primary outcome, the mean percentage of LEJR admissions

Table 1. Characteristics of the Study Population (All Eligible MSAs)<sup>a</sup>

Characteristics	Mean (SD)	Median (IQR)
Characteristics of eligible LEJR patients		
Age, y	72.5 (0.91)	72.5 (71.8-73.0)
White, %	90.0 (0.07)	92.2 (86.3-95.1)
Female, %	64.8 (0.03)	64.9 (62.7-67.0)
Disabled, %	17.2 (0.06)	16.0 (13.1-20.2)
Also receiving Medicaid, %	11.7 (0.06)	10.8 (7.9-14.3)
Elixhauser Comorbidity Index <sup>b</sup>	2.4 (0.29)	2.4 (2.2-2.5)
Characteristics of eligible hospitals		
CJR-eligible LEJR volume per hospital per month, No. <sup>c</sup>	20.4 (11.6)	17.5 (12.8-23.8)
For-profit hospitals, % <sup>d</sup>	25.3 (25.8)	25.0 (0-50.0)
Teaching hospitals, % <sup>d</sup>	9.3 (16.1)	0 (0-13.7)
Hospital beds, No. <sup>d</sup>	289.1 (140.9)	258.1 (205.0-332.8)
Characteristics of eligible MSAs		
Acute care hospitals, No. <sup>e</sup>	14.8 (16.0)	9.0 (6.0-18.0)
Institutional postacute care facilities, No. <sup>e</sup>	50.9 (53.3)	32.0 (20.0-58.0)
Home health agencies, No. <sup>e</sup>	57.8 (117.1)	24.0 (13.0-59.0)

Abbreviations: CJR, Comprehensive Care for Joint Replacement program; IQR, interquartile range; LEJR, lower extremity joint replacement; MSA, metropolitan statistical area.

<sup>a</sup> Table reports characteristics of CJR-eligible patients with Medicare who underwent LEJR (n=131 285 episodes), CJR-eligible hospitals (n=1633 hospitals), and CJR-eligible MSAs (n=196 MSAs). All data are reported at the MSA level. Means for characteristics of eligible LEJR patients and eligible hospitals are taken first within MSAs and then across MSAs. Unless otherwise noted, all data are based on 2016 Medicare claims data; episodes starting between April 1, 2016, and September 15, 2016, are included. At the patient level, the mean age was 72.6 (SD, 8.5) years (n=130 343) and the mean Elixhauser Comorbidity Index was 2.4 (SD, 1.7) (n=130 343).

<sup>b</sup> The Elixhauser Comorbidity Index is a comorbidity measure based on *International Classification of Diseases, Ninth Revision, Clinical Modification* or *International Statistical Classification of Diseases and Related Health Problems, Tenth Revision* diagnosis codes. Binary indicators were created for a list of 31 comorbidities; the index was created by summing these 31 indicators. An index of 0 means no comorbidities were present for the episode and a higher index means more comorbidities. The range of this measure is 0 to 31.

<sup>c</sup> Mean number of CJR-eligible LEJR episodes across hospitals in each MSA.

<sup>d</sup> Based on 2014 American Hospital Association data.

<sup>e</sup> Based on 2016 Medicare Provider of Services file.

discharged to institutional postacute care was 33.7% (SD, 11.2%) in the control group and was 2.9 percentage points lower (95% CI, -4.95 to -0.90 percentage points;  $P = .005$ ) in the CJR group, a significant difference.

Table 3 also shows secondary outcomes. The mean percentage of LEJR patients discharged to home without home health care was 32.2% (SD, 23.3%) in the control group and was 2.6 percentage points higher (95% CI, -0.79 to 5.90 percentage points;  $P = .14$ ) in the CJR group, a statistically non-significant difference. There was no statistically significant relationship between inclusion in CJR and discharges to other destinations.

Medicare spending on institutional postacute care was \$3871 (SD, \$1394) in the control group and was \$307 lower (95% CI, -\$587 to -\$27;  $P = .04$ ) in the CJR group, a statisti-

cally significant difference. There was no statistically significant relationship between inclusion in CJR and total Medicare FFS spending per episode, either gross or net of CJR reconciliation payments. Mean total Medicare spending per LEJR episode was \$22 872 (SD, \$3619) in the control group. Excluding reconciliation payments, total Medicare spending per LEJR episode was \$453 lower (95% CI, -\$909 to \$3;  $P = .06$ ) in the CJR group, and inclusive of reconciliation payments, total Medicare spending per LEJR episode was \$234 higher (95% CI, -\$214 to \$683;  $P = .31$ ) in the CJR group, a non-significant difference.

### Health Care Quality and Volume

Table 4 shows no statistically significant or substantively meaningful relationship between inclusion in CJR and any of the targeted quality or nontargeted quality measures. For example, the 90-day emergency department visit rate was 20.1% (SD, 2.9%) for the control group, and was 0.25 percentage points higher (95% CI, -0.44 to 0.93 percentage points;  $P = .48$ ) in the CJR group, a statistically nonsignificant difference.

Table 4 also shows no statistically significant relationship between inclusion in CJR and patient volume or case mix. For example, the mean number of CJR-eligible admissions per 1000 enrollees was 7.2 (SD, 3.5) in the control group and was 0.05 higher (95% CI, -0.32 to 0.42;  $P = .80$ ) in the CJR group; the mean Elixhauser Comorbidity Index was 2.3 (SD, 0.27) in the control group and was 0.01 lower (95% CI, -0.07 to 0.05;  $P = .73$ ) in the CJR group. eTables 6 and 7 in Supplement 1 further show no statistically significant relationship between inclusion in CJR and patient admissions under the various exclusion criteria (such as readmission for LEJR or death) and other measures of case mix (such as age or number of Charlson comorbidities<sup>27,30</sup>).

eTable 8 in Supplement 1 shows that not controlling for lags of the dependent variable reduced precision but did not substantively change the results. eTables 9 and 10 in Supplement 1 show intention-to-treat analysis of the relationship between original assignment to CJR and the outcomes, and eTable 11 and 12 in Supplement 1 show raw means for the outcome data by strata.

## Discussion

In an instrumental variable analysis of the first year of the CJR bundled payment model, MSAs that were covered by the CJR model, compared with those that were not, had a significantly lower percentage of patient discharges to institutional postacute care and significantly lower spending on institutional postacute care but no significant change in overall Medicare expenditures, particularly after accounting for the reconciliation (ie, bonus) payments. These results are consistent with previous work suggesting that reductions in postacute care are the first-line response of health systems to the introduction of alternative payment mechanisms.<sup>31</sup> There was no significant relationship between inclusion in CJR and targeted or nontargeted measures of quality of care, nor did hospitals appear to change

Table 2. Balance of Study Population by Group Prior to Implementation of CJR (2014)<sup>a</sup>

	Mean (SD) Value in Control MSAs (n = 121)	Mean Difference (95% CI) Between Original Bundled Payment and Control MSAs <sup>b</sup>	P Value <sup>c</sup>
Discharge destination, %			
Institutional postacute care <sup>d</sup>	41.0 (13.1)	-0.4 (-1.7 to 0.84)	.51
Home health agency <sup>e</sup>	31.3 (17.0)	1.4 (-0.08 to 2.9)	.07
Home (without home health agency) <sup>e</sup>	25.1 (21.6)	-0.7 (-2.2 to 0.67)	.31
Other <sup>e,f</sup>	2.5 (4.4)	-0.03 (-0.41 to 0.34)	.88
No. of days in institutional postacute care <sup>g</sup>	10.4 (2.9)	0.04 (-0.41 to 0.48)	.88
Medicare episode spending in institutional postacute care, \$	4719 (1410)	24 (-136 to 185)	.77
Total Medicare episode spending, \$	24 160 (3465)	-29 (-300 to 243)	.84
90-d Complication rate, % <sup>h</sup>	1.4 (0.69)	0.01 (-0.18 to 0.21)	.89
90-d Emergency department visits during episode, %	20.1 (3.1)	0.56 (-0.20 to 1.3)	.15
90-d All-cause readmission rate, % <sup>i</sup>	10.9 (2.2)	-0.08 (-0.69 to 0.54)	.81
LEJR admissions, No. per 1000 enrollees	8.3 (4.3)	0.13 (-0.09 to 0.34)	.26
CJR-eligible LEJR admissions, No. per 1000 enrollees <sup>j</sup>	6.9 (3.2)	0.06 (-0.23 to 0.36)	.69
Elixhauser Comorbidity Index <sup>e,k</sup>	2.4 (0.28)	0.01 (-0.04 to 0.09)	.75

Abbreviations: CJR, Comprehensive Care for Joint Replacement program; LEJR, lower extremity joint replacement; MSA, metropolitan statistical area.

<sup>a</sup> Table reports MSA-level average characteristics of CJR-eligible LEJR episodes in 2014. An *F* test of the hypothesis that all of the outcomes are jointly zero yielded an *F* statistic of 1.05 and *P* = .54. *N* = 196 MSAs included overall. All measures are based on 2014 Medicare claims data; episodes starting between April 1, 2014, and September 15, 2014, are included.

<sup>b</sup> Estimated difference between the 121 control MSAs and the original 75 treatment MSAs from an ordinary least squares regression of the outcome on an indicator variable for original treatment status and controls for strata and for 2 lags of the dependent variable (measured in 2012 and 2013).

<sup>c</sup> *P* values for differences are based on heteroskedasticity-robust standard errors.

<sup>d</sup> Institutional postacute care includes skilled nursing facilities, long-term care hospitals, and inpatient rehabilitation facilities.

<sup>e</sup> Outcome not prespecified in analysis plan.

<sup>f</sup> Includes Medicare-approved swing beds and inpatient care at other acute care hospitals, intermediate care facilities, and other less common destinations such as psychiatric hospitals, hospice care, and federal hospitals.

<sup>g</sup> Sum of days in skilled nursing facilities, long-term care hospitals, and inpatient

rehabilitation facilities. This measure is averaged across all episodes, not just episodes with institutional postacute care use.

<sup>h</sup> Complication rate is the share of CJR-eligible patients who have at least 1 of the 8 underlying complications that contribute to the total hip arthroplasty/total knee arthroplasty 90-day complication measure used in the targeted quality score; unlike the input into the targeted quality score, the complication measure here is not risk-standardized and is measured for admissions during the study period rather than as a 3-year moving average. The sample for this measure includes patients who were readmitted for LEJR and those who died within the episode, both of which would otherwise be excluded from the CJR sample based on patient eligibility criteria.

<sup>i</sup> The sample for this measure includes patients who were readmitted for LEJR, who would otherwise be excluded from the CJR sample based on patient eligibility criteria.

<sup>j</sup> Some hospitals and some LEJR patients are excluded from CJR. All other analyses in the article are limited to CJR-eligible admissions unless otherwise specified.

<sup>k</sup> See footnote to Table 1 for detailed description.

their rates of admissions for covered patients or strategically admit patients with lower illness severity, as some observers have feared.<sup>2</sup>

The relationship between CJR and health care use and spending was smaller than those from prior observational studies of Medicare bundled payment programs. A difference-in-differences analysis of the associations of BPCI for LEJR found a 4% decline in Medicare spending<sup>12</sup>; a pre-post analysis of BPCI for LEJR in one health system found a 21% decline and received considerable attention in the media.<sup>11,32</sup> In addition, a matched-control study of Medicare's Heart Bypass Center Demonstration Project estimated net savings of 14% per episode<sup>8</sup> and was cited as reason to expect substantial savings from subsequent bundled payment programs.<sup>3</sup>

There are several possible reasons the findings from the current study of a randomized, mandatory bundled payment program contrast with prior studies of voluntary bundled payment models for LEJR. One reason is selection on expected costs. Estimates of the savings from voluntary programs may be biased upward if hospitals with lower expected spending

are more likely to sign up; fully controlling for this can be challenging.<sup>13</sup> A second reason is selection on treatment effects.<sup>33</sup> The effect of bundled payment on the subsample of hospitals that select into the voluntary program may not be representative of the average effect across all hospitals.

A third reason is the size of the incentives. In the first year of CJR, financial incentives were about one-fourth the size of those in the voluntary BPCI model and had no downside risk; by years 4 and 5, CJR incentives will be the same as those in BPCI. An examination of the effects of CJR in subsequent years will be important, as the phase-in of downside risk and larger incentives may reduce Medicare spending (holding hospital behavior fixed) and may cause larger changes in hospital behavior. However, a comprehensive analysis of the CJR program as initially designed will not be possible. In December 2017, CMS modified CJR to be voluntary for 33 of the 67 included MSAs starting in performance year 3. This means that starting in performance year 3, the treatment vs control analysis performed herein can only be implemented for the remaining 34 MSAs where enrollment is mandatory.<sup>34</sup>

Table 3. Health Care Use and Spending During First Year of CJR (2016)<sup>a</sup>

	Mean (SD) Value in Control MSAs (n = 121)	Mean Difference (95% CI) Between Bundled Payment and Control MSAs	P Value <sup>b</sup>
Discharge destination, %			
Institutional postacute care <sup>c</sup>	33.7 (11.2)	-2.9 (-4.95 to -0.90)	.005
Home health agency <sup>d</sup>	32.2 (19.4)	1.3 (-1.73 to 4.34)	.40
Home (without home health agency) <sup>d</sup>	32.2 (23.3)	2.6 (-0.79 to 5.90)	.14
Other <sup>d</sup>	1.9 (3.0)	-0.40 (-0.84 to 0.05)	.08
No. of days in institutional postacute care	8.2 (2.7)	-0.48 (-1.06 to 0.11)	.12
Medicare episode spending in institutional postacute care, \$	3871 (1394)	-307 (-587 to -27)	.04
Total Medicare episode spending, \$ <sup>e</sup>	22 872 (3619)	-453 (-909 to 3)	.06
Medicare spending net of reconciliation payment, \$ <sup>e,f</sup>	22 872 (3619)	234 (-214 to 683)	.31

Abbreviations: CJR, Comprehensive Care for Joint Replacement program; MSA, metropolitan statistical area.

<sup>a</sup> Change with bundled payment was estimated with the use of 2-stage least-squares instrumental variable regression, with random assignment to CJR used as an instrument for inclusion in CJR. All regressions include indicators for strata and control for 2 lags of the dependent variable. All outcomes were measured during the episode. All measures are based on 2016 Medicare claims data unless otherwise noted; episodes starting between April 1, 2016, and September 15, 2016, are included; lags of the dependent variable are based on the corresponding periods in 2013 and 2014. N = 196 MSAs included overall. See Table 2 for variable definitions.

<sup>b</sup> P values for differences are based on heteroskedasticity-robust standard errors.

<sup>c</sup> Primary outcome.

<sup>d</sup> The preanalysis plan specified that discharges to home health agencies would be combined with discharges to institutional postacute care rather than being analyzed separately, and analysis of the remaining categories was not specified. The changes are nonsubstantive and designed for ease of exposition.

<sup>e</sup> Because control MSAs do not receive reconciliation payments, control MSA total Medicare episode spending and Medicare spending net of reconciliation payment are identical.

<sup>f</sup> Medicare spending net of reconciliation payment equals total Medicare episode spending (previous row) plus any reconciliation payments made by Medicare. Data on reconciliation payment were downloaded from <https://innovation.cms.gov/Files/x/cjr-py1reconpym.xlsx>. Because reconciliation payment did not exist prior to 2016, analyses instead controlled for lags of total Medicare episode spending.

Table 4. Quality and Patient Case Mix During First Year of CJR (2016)<sup>a</sup>

	Mean (SD) Value in Control MSAs (n = 121)	Mean Difference (95% CI) Between Bundled Payment and Control MSAs	P Value <sup>b</sup>
Quality measures			
Targeted measure: composite quality score <sup>c</sup>	9.3 (2.8)	0.72 (-0.12 to 1.55)	.10
90-d Nontargeted quality measures, %			
Complication rate	1.2 (0.71)	-0.002 (-0.21 to 0.21)	.99
Emergency department visit rate	20.1 (2.9)	0.25 (-0.44 to 0.93)	.48
All-cause readmission rate	10.1 (2.0)	0.01 (-0.55 to 0.58)	.97
Admissions and patient case mix			
LEJR admissions, No. per 1000 enrollees	9.0 (4.8)	-0.14 (-0.42 to 0.14)	.34
CJR-eligible LEJR admissions, No. per 1000 enrollees	7.2 (3.5)	0.05 (-0.32 to 0.42)	.80
Elixhauser Comorbidity Index <sup>d</sup>	2.3 (0.27)	-0.01 (-0.07 to 0.05)	.73

Abbreviations: CJR, Comprehensive Care for Joint Replacement program; LEJR, lower extremity joint replacement; MSA, metropolitan statistical area.

<sup>a</sup> Change with bundled payment was estimated with the use of 2-stage least squares instrumental variable regression with random assignment to CJR used as an instrument for inclusion in CJR. All regressions include indicators for strata and control for 2 lags of the dependent variable. All outcomes are based on 2016 Medicare claims data unless otherwise noted; episodes starting between April 1, 2016, and September 15, 2016, are included. N = 196 MSAs included overall.

<sup>b</sup> P values for differences are based on heteroskedasticity-robust standard errors.

<sup>c</sup> This measure ranges from 0 to 18, with higher scores indicating higher quality. The instrumental variable estimate does not include lags of the dependent variable because it involves 3-year moving averages. As discussed in the text, this measure was not expected to be affected because the score for the first year was based almost entirely on data prior to the introduction of CJR.

<sup>d</sup> See footnote to Table 1 for detailed description. Outcome not prespecified in analysis plan.

**Limitations**

This study has several limitations. First, this study evaluated outcomes only during the first year following implementation of the bundled payment program, when the maximum stop-gain and stop-loss incentives had not yet been implemented. Second, the analysis was powered to detect differences in discharges to institutional postacute care but not

the corresponding difference in total episode Medicare spending. Third, the analysis did not explore heterogeneity in effects of bundled payments across patients or hospitals, which might shed light on potential mechanisms; in this spirit, early work has compared characteristics of hospitals that did and did not achieve shared savings in the first year.<sup>35</sup> Fourth, potentially important health outcomes such as pain

or functional limitations were not analyzed. Fifth, analysis of the relationship between bundled payments and the targeted quality measure is not yet meaningful, given that in the first year it was almost entirely measured prior to the start of CJR. Sixth, the analysis pertained only to health care use in the bundle during the covered episode; as more data become available, it will be of interest to see what happens to health care use and spending over longer horizons and health care use outside of the bundle (such as prescription pain medication covered by Medicare Part D).

## Conclusions

In this interim analysis of the first year of the CJR bundled payment model for LEJR among Medicare beneficiaries, MSAs covered by CJR, compared with those that were not, had a significantly lower percentage of discharges to institutional postacute care but no significant difference in total Medicare spending per LEJR episode. Further evaluation is needed as the program is more fully implemented.

### ARTICLE INFORMATION

**Accepted for Publication:** August 2, 2018.

**Author Contributions:** Ms Ji and Dr Skinner had full access to all the data in the study and take responsibility for the integrity of the data and the accuracy of the data analysis.

**Concept and design:** Finkelstein, Mahoney.

**Acquisition, analysis, or interpretation of data:** All authors.

**Drafting of the manuscript:** Finkelstein, Mahoney.  
**Critical revision of the manuscript for important intellectual content:** All authors.

**Statistical analysis:** All authors.

**Obtained funding:** Finkelstein, Mahoney, Skinner.  
**Administrative, technical, or material support:** All authors.

**Supervision:** Finkelstein, Mahoney.

**Conflict of Interest Disclosures:** All authors have completed and submitted the ICMJE Form for Disclosure of Potential Conflicts of Interest. Dr Skinner reports that he is an investor in Dorsata Inc, a clinical pathway software startup, and a consultant to Sutter Health Inc. No other disclosures were reported.

**Funding/Support:** J-PAL North America and the National Institute on Aging (grant P01AG019783-15) provided research support.

**Role of the Funder/Sponsor:** The funders had no role in the design and conduct of the study; collection, management, analysis, and interpretation of the data; preparation, review, or approval of the manuscript; or decision to submit the manuscript for publication.

**Data Sharing Statement:** See Supplement 2.

### REFERENCES

- Shatto JD. Center for Medicare and Medicaid Innovation's methodology and calculations for the 2016 estimate of fee-for-service payments to alternative payment models. March 3, 2016. <https://innovation.cms.gov/Files/x/ffs-apm-goalmemo.pdf>. Accessed July 18, 2018.
- Fisher ES. Medicare's bundled payment program for joint replacement: promise and peril? *JAMA*. 2016;316(12):1262-1264. doi:10.1001/jama.2016.12525
- Cutler D. How health care reform must bend the cost curve. *Health Aff (Millwood)*. 2010;29(6):1131-1135. doi:10.1377/hlthaff.2010.0416
- Cutler DM, Ghosh K. The potential for cost savings through bundled episode payments. *N Engl J Med*. 2012;366(12):1075-1077. doi:10.1056/NEJMp113361
- Furman J, Kocher B. A health-care fix that works, now being rolled back. *Wall Street Journal*. August 20, 2017. <https://www.wsj.com/articles/a-health-care-fix-that-works-now-being-rolled-back-1503258369>. Accessed March 14, 2018.
- Mechanic RE. Mandatory Medicare bundled payment—is it ready for prime time? *N Engl J Med*. 2015;373(14):1291-1293. doi:10.1056/NEJMp1509155
- Miller DC, Gust C, Dimick JB, Birkmeyer N, Skinner J, Birkmeyer JD. Large variations in Medicare payments for surgery highlight savings potential from bundled payment programs. *Health Aff (Millwood)*. 2011;30(11):2107-2115. doi:10.1377/hlthaff.2011.0783
- 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.
- Doran JP, Zabinski SJ. Bundled payment initiatives for Medicare and non-Medicare total joint arthroplasty patients at a community hospital: bundles in the real world. *J Arthroplasty*. 2015;30(3):353-355. doi:10.1016/j.arth.2015.01.035
- Froemke CC, Wang L, DeHart ML, Williamson RK, Ko LM, Duweliu PJ. Standardizing care and improving quality under a bundled payment initiative for total joint arthroplasty. *J Arthroplasty*. 2015;30(10):1676-1682. doi:10.1016/j.arth.2015.04.028
- Navathe AS, Troxel AB, Liao JM, et al. Cost of joint replacement using bundled payment models. *JAMA Intern Med*. 2017;177(2):214-222. doi:10.1001/jamainternmed.2016.8263
- 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. doi:10.1001/jama.2016.12717
- Gronniger T, Fiedler M, Patel K, Adler L, Ginsberg P. How should the Trump Administration handle Medicare's new bundled payment programs? *Health Affairs* blog. April 2017. <https://www.brookings.edu/blog/usc-brookings-schaeffer-on-health-policy/2017/04/10/how-should-the-trump-administration-handle-medicare-new-bundled-payment-programs/>. Accessed July 18, 2018.
- Centers for Medicare & Medicaid Services. National Summary of Inpatient Charge Data by Medicare Severity Diagnosis Related Group (MS-DRG), FY2014. 2014. <https://data.cms.gov/Medicare-Inpatient/National-Summary-of-Inpatient-Charge-Data-by-Medic/sfua-yggc>. Accessed June 5, 2018.
- US Department of Health and Human Services. 2014 CMS Statistics. 2014. [https://www.cms.gov/Research-Statistics-Data-and-Systems/Statistics-Trends-and-Reports/CMS-Statistics-Reference-Booklet/Downloads/CMS\\_Stats\\_2014\\_final.pdf](https://www.cms.gov/Research-Statistics-Data-and-Systems/Statistics-Trends-and-Reports/CMS-Statistics-Reference-Booklet/Downloads/CMS_Stats_2014_final.pdf). Accessed June 5, 2018.
- Centers for Medicare & Medicaid Services. Medicare program; hospital inpatient prospective payment systems for acute care hospitals and the long-term care hospital prospective payment system policy changes and fiscal year 2016 rates; revisions of quality reporting requirements for specific providers, including changes related to the electronic health record incentive program; extensions of the Medicare-dependent, small rural hospital program and the low-volume payment adjustment for hospitals: final rule; interim final rule with comment period. *Fed Regist*. 2015;80(158):49325-49886.
- Finkelstein A, Taubman S, Wright B, et al; Oregon Health Study Group. The Oregon Health Insurance Experiment: evidence from the first year. *Q J Econ*. 2012;127(3):1057-1106. doi:10.1093/qje/qjs020
- Baicker K, Taubman SL, Allen HL, et al; Oregon Health Study Group. The Oregon experiment—effects of Medicaid on clinical outcomes. *N Engl J Med*. 2013;368(18):1713-1722. doi:10.1056/NEJMsal121321
- Centers for Medicare & Medicaid Services. Medicare program; comprehensive care for joint replacement payment model for acute care hospitals furnishing lower extremity joint replacement services. *Fed Regist*. 2015;80(226):73273-73554.
- Centers for Medicare & Medicaid Services. *Overview of CJR Measures, Composite Quality Score, and Pay-For-Performance Methodology*. 2018. <https://innovation.cms.gov/Files/x/cjr-qualsup.pdf>. Accessed July 18, 2018.
- Baicker K, Finkelstein A, Song J, Taubman S. The impact of Medicaid on labor market activity and program participation: evidence from the Oregon Health Insurance Experiment. *Am Econ Rev*. 2014;104(5):322-328. doi:10.1257/aer.104.5.322
- Taubman SL, Allen HL, Wright BJ, Baicker K, Finkelstein AN. Medicaid increases emergency-department use: evidence from Oregon's Health Insurance Experiment. *Science*. 2014;343(6168):263-268. doi:10.1126/science.1246183
- Finkelstein AN, Taubman SL, Allen HL, Wright BJ, Baicker K. Effect of Medicaid coverage on ED use—further evidence from Oregon's experiment. *N Engl J Med*. 2016;375(16):1505-1507. doi:10.1056/NEJMp1609533
- Centers for Medicare and Medicaid Services. Medicare program; hospital inpatient prospective



payment systems for acute care hospitals and the long-term care hospital prospective payment system and policy changes and fiscal year 2017 rates; quality reporting requirements for specific providers; graduate medical education; hospital notification procedures applicable to beneficiaries receiving observation services; technical changes relating to costs to organizations and Medicare cost reports; finalization of interim final rules with comment period on LTCH PPS payments for severe wounds, modifications of limitations on redesignation by the Medicare Geographic Classification Review Board, and extensions of payments to MDHs and low-volume hospitals: final rule. *Fed Regist*. 2016;81(162):56761-57345.

25. Centers for Medicare & Medicaid Services. Hospital Compare datasets. <https://data.medicare.gov/data/hospital-compare>. Accessed October 4, 2017.

26. Elixhauser A, Steiner C, Harris DR, Coffey RM. Comorbidity measures for use with administrative data. *Med Care*. 1998;36(1):8-27. doi:10.1097/00005650-199801000-00004

27. Quan H, Sundararajan V, Halfon P, et al. Coding algorithms for defining comorbidities in ICD-9-CM and ICD-10 administrative data. *Med Care*. 2005;43(11):1130-1139. doi:10.1097/01.mlr.0000182534.19832.83

28. Colla C, Bynum J, Austin A, Skinner J. *Hospital Competition, Quality, and Expenditures in the US Medicare Population*. Cambridge, MA: National Bureau of Economic Research; November 2016. NBER working paper 22826.

29. Imbens GW, Angrist JD. Identification and estimation of local average treatment effects. *Econometrica*. 1994;62(2):467-475. doi:10.2307/2951620

30. Charlson ME, Pompei P, Ales KL, MacKenzie CR. A new method of classifying prognostic comorbidity in longitudinal studies: development and validation. *J Chronic Dis*. 1987;40(5):373-383. doi:10.1016/0021-9681(87)90171-8

31. 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(4):518-526. doi:10.1001/jamainternmed.2016.9115

32. Meyer H. Bundled-payment joint replacement programs winning over surgeons. *Modern Healthcare*. October 2017. <http://www.modernhealthcare.com/article/20171007/NEWS/171009950>. Accessed July 18, 2018.

33. Heckman JJ, Vytlacil E, Urzua S. Understanding instrumental variables in models with essential heterogeneity. *Rev Econ Stat*. 2006;LXXXVIII(3). doi:10.1162/rest.88.3.389

34. Centers for Medicare & Medicaid Services. Comprehensive Care for Joint Replacement model. 2018. <https://innovation.cms.gov/initiatives/cjr>. Accessed July 18, 2018.

35. 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(9):930-932. doi:10.1001/jama.2018.0678