Original Investigation

Association Between a Bundled Payment Program for Lower Extremity Joint Replacement and Patient Outcomes Among Medicare Advantage Beneficiaries

Amanda Sutherland, PhD; Emily Boudreau, PhD; Andy Rowe, MPH; Qian Huang, MPH; Joshua M. Liao, MD, MSc; Matthew Flagg, MBA; Deborah Cousins, MSPH; Dana Drzayich Antol, MS; William H. Shrank, MD, MSHS; Brian W. Powers, MD, MBA; Amol S. Navathe, MD, PhD

Abstract

IMPORTANCE Much of the evidence for bundled payments has been drawn from models in the traditional Medicare program. Although private insurers are increasingly offering bundled payment programs, it is not known whether they are associated with changes in episode spending and quality.

OBJECTIVE To evaluate whether a voluntary bundled payment program offered by a national Medicare Advantage insurer was associated with changes in episode spending or quality of care for beneficiaries receiving lower extremity joint replacement (LEJR) surgery.

DESIGN, SETTING, AND PARTICIPANTS Cross-sectional study of 23,034 LEJR surgical episodes that emulated a stepped-wedge design by using the time-varying, geographically staggered rollout of the bundled payment program from January 1, 2012, to September 30, 2019. Episode-level multivariable regression models were estimated within practice to compare changes before and after program participation, using episodes at physician practices that had not yet begun participating in the program during a given time period (but would go on to do so) as the control. Data analyses were performed from July 1, 2021, to June 30, 2022.

EXPOSURES Physician practice participation in the bundled payment program.

MAIN OUTCOMES AND MEASURES The primary outcome was episode spending (plan and beneficiary). Secondary outcomes included postacute care use (skilled nursing facility and home health care), surgical setting (inpatient vs outpatient), and quality (90-day complications [including deep vein thrombosis, wound infection, fracture, or dislocation] and readmissions).

RESULTS The final analytic sample included 23,034 LEJR episodes (6,355 bundled episodes and 16,679 control episodes) from 109 physician practices participating in the program. Of the beneficiaries, 7,730 were male and 15,304 were female; 3,057 were Black, 19,351 were White, 447 were of other race or ethnicity (assessed according to the Centers for Medicare & Medicaid Services beneficiary race and ethnicity code, which reflects data reported to the Social Security Administration), and 179 were of unknown race and ethnicity. The mean (SD) age was 70.9 (7.2) years. Participation in the bundled payment program was associated with a 2.7% (95% CI, 1.3%-4.1%) decrease in spending per episode (mean episodic spending, $21,964 [95% CI, $21,636-$22,296] vs $22,562 [95% CI, $22,346-$22,779]), as well as reductions in skilled nursing facility use after discharge (21.3% for bundled episodes vs 25.0% for control episodes; odds ratio [OR], 0.81 [95% CI, 0.67-0.98]) and increased use of the outpatient surgical setting (14.1% for bundled episodes vs 8.4% for control episodes; OR, 1.79 [95% CI, 1.53-2.09]). The program was not associated with changes in quality outcomes, including 90-day complications (8.8% for bundled episodes vs 8.6% for control episodes).

Key Points

Question Was there an association between a bundled payment program and changes in total episodic spending or quality of care for Medicare Advantage beneficiaries receiving lower extremity joint replacement surgery?

Findings In this cross-sectional study of 23,034 lower extremity joint replacement surgical episodes, physician practice participation in a bundled payment program was associated with a 2.7% reduction in episode spending without changes in quality.

Meaning This study found that bundled payments offered by private insurers, including Medicare Advantage plans, may have the ability to reduce spending for lower extremity joint replacement episodes while maintaining quality of care.

+ Supplemental content

Author affiliations and article information are listed at the end of this article.

Open Access. This is an open access article distributed under the terms of the CC-BY-NC-ND License.
episodes; OR, 1.02 [95% CI, 0.86-1.20]) and readmissions (4.3% for bundled episodes vs 4.6% for control episodes; OR, 0.92 [95% CI, 0.75-1.13]).

CONCLUSIONS AND RELEVANCE  In this study of an LEJR bundled payment program offered by a national Medicare Advantage insurer, findings suggest that physician practice participation in the program was associated with a decrease in episode spending without changes in quality. Bundled payments offered by private insurers, including Medicare Advantage plans, are an alternate payment option to fee for service that may reduce spending for LEJR episodes while maintaining quality of care.


Introduction

During the last decade, private and public insurers in the US have tested payment approaches to reduce unwarranted variation in spending, use, and quality. One strategy has been bundled payments, a value-based payment arrangement with promising evidence, particularly for surgical episodes. Under bundled payments, physicians and health care organizations receive a predetermined payment for an entire episode of surgical care, including preoperative and postoperative care.

The Centers for Medicare & Medicaid Services (CMS) has implemented a number of bundled payment programs for hip and knee replacement (hereafter, lower extremity joint replacement [LEJR]) within the traditional Medicare program, which include the voluntary Bundled Payments for Care Improvement and Bundled Payments for Care Improvement Advanced initiatives and the mandatory Comprehensive Care for Joint Replacement program. These programs have been associated with lower episodic spending and stable or improved quality of care for patients with LEJR.1-3

Private insurers are increasingly offering bundled payments for LEJR, but there is a lack of data on the repercussions of these programs.4-6 This gap is particularly notable in Medicare Advantage because approximately half of all Medicare beneficiaries are now choosing to enroll in Medicare Advantage.7,8

In this study, we examined the association between physician practice participation in a bundled payment program for LEJR offered by a large national Medicare Advantage insurer and changes in episode spending, surgical setting, postacute care use, and quality.

Methods

Overview of Program

This cross-sectional study evaluated an LEJR bundled payment program offered by Humana Inc, a large national Medicare Advantage insurer. The program is ongoing, retrospective, and voluntary and offers participating physician practices shared savings for reducing spending and meeting clinical quality measures for LEJR episodes. From a program operations perspective, episodes include all spending from 45 days before LEJR surgery until 90 days afterward, with the exclusion of home health care and pharmacy spending. This program was designed to not penalize or dissuade use of certain services, such as home health care (however, in this study, spending on these categories was reinstated). Episodic spending targets are set for each physician practice according to historical LEJR spending. The spending targets do not include any adjustment for treatment setting, so shifting patients from inpatient to outpatient surgical settings was a deliberate mechanism of savings. To receive shared savings, physician practices must also meet quality standards for surgical
complications and readmissions. Additional information on the program is available in the eAppendix in Supplement 1.

Two main features differentiate this program from bundled payment programs in traditional Medicare: (1) the program includes LEJR episodes occurring in inpatient or outpatient settings, whereas programs in traditional Medicare have mostly focused on inpatient episodes; and (2) only physician practices can participate in the program, whereas programs in traditional Medicare allow participation from physician practices and hospitals and have historically engaged more hospitals than physician practices.

Physician practices were eligible to join the program on a rolling basis by geographic area, starting in January 2016. Data on physician practice eligibility by geography are available in eTable 1 in Supplement 1.

**Study Design**

We used the time-varying, geographically staggered rollout of the program to emulate a stepped-wedge study design from January 1, 2012, to December 31, 2019. The sequencing of the rollout was based on operational considerations and not historical cost or quality outcomes of physician practices in the region, creating a plausibly exogenous source of variation for this study. We estimated within-practice changes before and after program participation, using episodes at physician practices that had not yet begun participating in the program during a given time period (but would go on to do so) as the control. This approach, which has been used previously to study private insurer payment programs with staggered rollouts, mitigated bias from physician practice selection into the program on unobserved factors because all practices in our sample eventually elected to participate. It also allowed us to account for secular trends during the study period.

**Data Sources**

We used program data from the Medicare Advantage insurer to determine when physician practices began participating in the program. To construct our outcomes, we used longitudinal medical claims, allowed amounts, and beneficiary eligibility information from the Medicare Advantage insurer.

**Study Period and Episode Identification**

Our primary unit of analysis was the LEJR episode. All LEJR episodes included in the analysis were indexed around a surgical date that occurred during the study period from January 1, 2012, to September 30, 2019. The latter date was chosen to allow for 90 days of postsurgery follow-up in 2019, which was necessary to construct our outcome measures. We did not extend our analysis into 2020 because of the COVID-19 pandemic.

To identify eligible LEJR episodes, we started by identifying the 109 physician practices that participated in the program at any point from January 1, 2016, to December 31, 2020. Extending the window of physician practice identification beyond the study period (which ended September 30, 2019) allowed us to identify additional control episodes for 2019 (ie, episodes performed by physician practices that were not participating in 2019 but went on to do so in 2020). Next, we identified all enrolled beneficiaries receiving LEJR surgery by a participating practice. We selected 1 procedure at random per beneficiary.

We excluded episodes for beneficiaries not continuously enrolled in the Medicare Advantage plan for at least 6 months before and 3 months after the surgery, enrolled in hospice during the study period, dually enrolled in Medicare and Medicaid, who left the hospital against medical advice after surgery, who were attributed to an organization that delegated claims processing to a third party, and who were contractually excluded from research. To align with the operational exclusions used in this specific program, we also excluded episodes for beneficiaries with specific high-risk, high-cost conditions (eg, trauma or cancer), as well as those with insufficient claims, suggesting incomplete or inconsistent billing. A list of clinical and program-related exclusions is available in the eAppendix in Supplement 1.
Assignment of Exposure Status

The LEJR episodes that occurred before the physician practice began participating in the bundled payment program were considered control episodes. Surgical procedures occurring at least 6 months after the practice began participating were considered bundled episodes. Consistent with prior research, a 6-month washout period was included to allow for practice-level changes initiated by newly participating physician practices in response to program incentives. Episodes performed during the washout period were excluded from the analysis. See Figure 1 for the distribution of bundled episodes during the study period.

Outcomes

Our primary outcome was episode spending, which included all plan and beneficiary spending from 45 days before surgery through 90 days after facility discharge. All amounts were adjusted to 2019 dollars with the medical Consumer Price Index.

Secondary outcomes included postacute care use, surgical setting, and measures of episode quality. To evaluate postacute care use in the 90 days after facility discharge, we identified skilled nursing facility (SNF) use and home health care use that were initiated in the first 10 days after the surgery as separate, binary outcomes. Although home health care spending was excluded from the spending benchmark in the bundled payment program, we conducted an analysis on the association between program participation and home health care use. We did not examine inpatient rehabilitation facility or long-term acute care hospital use owing to low rates of use. We evaluated surgical setting by identifying whether the surgery was conducted in the inpatient or outpatient setting. We evaluated episode quality by assessing 90-day complications (a composite measure of any deep vein thrombosis, wound infection, fracture, or dislocation) and 90-day readmissions. These complications were included because physician practices are held accountable for them under the bundled payment program and they align with CMS bundled payment program quality metrics.

Statistical Analysis

We compared beneficiary characteristics for bundled and control episodes by using standardized mean differences. With respect to our primary outcome, we constructed episode-level generalized linear regression models to examine the association between program participation and total episode spending, using a gamma distribution with a log-link function to account for skewness in spending distributions. This method exploited the time-varying rollout of the program to estimate within-practice changes before and after program participation. The plausibly exogenous (ie, unrelated to physician practice performance) staggered rollout mitigated bias from unobserved selection into program participation and allowed us to control for secular trends.
We also included a cohort of physician practices that entered the program in 2020, which contributed only control episodes because our analytic study period stopped in 2019. To examine this assumption of exogenous exposure timing, we compared preexposure spending for early participants with that for those who participated later and did not find evidence of obvious differences (eFigure 1 in Supplement 1).

In addition to time (quarter-year) fixed effects and physician practice fixed effects, these models also adjusted for the following covariates: beneficiary age (<65, 65-74, 75-84, and ≥85 years), sex, race and ethnicity, Elixhauser comorbidity score, and procedure type (knee vs hip replacement). All covariates were measured on the date of the beneficiary's surgery, except for the Elixhauser comorbidity score, which was assessed during the 6 months before surgery. Race and ethnicity were assessed according to the CMS beneficiary race and ethnicity code and categorized as Black, White, other, and unknown owing to inaccuracies in classification for non-Black and non-White beneficiaries. The "other" category included the following races and ethnicities: Asian, Hispanic, North American Native, and other. Robust SEs were clustered at the physician practice tax identification number level. More details on our regression specification are provided in the eAppendix in Supplement 1.

For secondary outcomes of SNF use and 90-day complications, and for exploratory outcome of home health care use, we used similarly specified logistic regression models to account for the binary outcomes. For the secondary outcome of 90-day readmissions, physician practice fixed effects were removed because of model convergence issues.

For the secondary outcome of surgical setting, we used a slightly modified approach. To account for CMS rule changes in 2018 and 2019, which allowed knee and hip replacements to be completed in the outpatient setting, we added an indicator variable for whether the episode occurred after the CMS rule change (2018 for total knee replacement and 2019 for total hip replacement) and an interaction term between this indicator variable and the physician practices' program participation status to estimate the association between surgical setting and program participation in the post-CMS rule change time periods. Year fixed effects were included instead of quarter-year and physician practice fixed effects, and clustering and procedure type (knee vs hip replacement) were removed because of model convergence issues.

We conducted several sensitivity analyses to test the robustness of our findings. First, we used an alternative beneficiary comorbidity specification, replacing a beneficiary's Elixhauser score with his or her CMS Hierarchical Conditions Category score. Second, although earlier evaluations have shown little evidence for beneficiary selection after bundled payment program participation, to mitigate concerns that physician practices might select healthier beneficiaries after participation in the program, we assessed whether beneficiary comorbidity, frailty, disability, or chronic obstructive pulmonary disease status changed before and after program participation. To implement this assessment, we repeated our main regression models with beneficiary Elixhauser score, frailty status, disability status, and chronic obstructive pulmonary disease status as the dependent variables. Finally, we used a variety of sensitivity analyses to test robustness, including stratifying by early and late program joiners, stratifying by joint type (knees vs hips), changing the study years to 2015 to 2019, and removing episodes from markets that participated in the Comprehensive Care for Joint Replacement bundled payment program to mitigate concerns that spillover effects from that program might be affecting spending in those markets.

All analyses were performed with SAS Enterprise Guide, version 8.2 from July 1, 2021, to June 30, 2022. An Advarra institutional review board approved and granted a waiver of informed consent for this study. The waiver was granted because the research involved no more than minimal risk, subjects for whom records were reviewed were not readily available to sign a consent, a possibility existed of creating additional risks to privacy by linking otherwise deidentified data with nominal identifiers to contact individuals to seek consent, and the waiver did not adversely affect the rights and welfare of the subjects. This study followed the Strengthening the Reporting of Observational
Studies in Epidemiology (STROBE) reporting guideline for cross-sectional studies. All hypothesis tests were 2 sided, with \( P < .05 \) indicating significance.

## Results

### Sample Characteristics

The final analytic sample included 23,034 LEJR episodes, with 6,355 bundled episodes and 16,679 control episodes (see eFigures 2 and 3 in Supplement 1 for the attrition diagram and distribution of episodes by exposure status during the study period, respectively). Of the beneficiaries, 7,730 were male and 15,304 were female, 3,057 were Black, 19,351 were White, 447 were of other race or ethnicity (assessed according to the CMS beneficiary race and ethnicity code, which reflects data reported to the Social Security Administration), and 179 were of unknown race and ethnicity. The mean (SD) age was 70.9 (7.2) years. Beneficiaries in bundled episodes were more likely to reside in the South (owing to earlier rollout of the bundled payment program in this region) and had a higher comorbidity burden than beneficiaries of control episodes (Table 1).\(^{19}\)

### Total Episode Spending

In unadjusted analyses, spending for bundled episodes was $18,640 compared with $21,795 for control episodes (Figure 2; eTable 2 in Supplement 1). In adjusted analyses, participation in the program was associated with a 2.7% (95% CI, 1.3%-4.1%) decrease in episode spending compared with control episodes, representing a change of −$598 per episode (mean episodic spending, Table 1).

<table>
<thead>
<tr>
<th>Characteristics</th>
<th>Bundled episodes (( n = 6,355 ))</th>
<th>Control episodes (( n = 16,679 ))</th>
<th>SMD(^{a})</th>
</tr>
</thead>
<tbody>
<tr>
<td>Age category in years, No. (%)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>&lt;65</td>
<td>867 (13.6)</td>
<td>2,380 (14.3)</td>
<td>0.018</td>
</tr>
<tr>
<td>65-74</td>
<td>3,378 (53.2)</td>
<td>9,359 (56.1)</td>
<td>0.059</td>
</tr>
<tr>
<td>75-84</td>
<td>1,933 (30.4)</td>
<td>4,561 (27.3)</td>
<td>0.068</td>
</tr>
<tr>
<td>≥85</td>
<td>177 (2.8)</td>
<td>379 (2.3)</td>
<td>0.033</td>
</tr>
<tr>
<td>Sex, No. (%)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Female</td>
<td>4,273 (67.2)</td>
<td>11,031 (66.1)</td>
<td>0.023</td>
</tr>
<tr>
<td>Male</td>
<td>2,082 (32.8)</td>
<td>5,648 (33.9)</td>
<td>0.023</td>
</tr>
<tr>
<td>Race and ethnicity, No. (%)(^{b})</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Black</td>
<td>849 (13.4)</td>
<td>2,208 (13.2)</td>
<td>0.004</td>
</tr>
<tr>
<td>White</td>
<td>5,343 (84.1)</td>
<td>14,008 (84.0)</td>
<td>0.002</td>
</tr>
<tr>
<td>Other(^{c})</td>
<td>94 (1.5)</td>
<td>353 (2.1)</td>
<td>0.048</td>
</tr>
<tr>
<td>Unknown</td>
<td>69 (1.1)</td>
<td>110 (0.7)</td>
<td>0.046</td>
</tr>
<tr>
<td>Geographic region, No. (%)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Northeast</td>
<td>19 (0.3)</td>
<td>209 (1.3)</td>
<td>0.109</td>
</tr>
<tr>
<td>Midwest</td>
<td>1,333 (21.0)</td>
<td>4,431 (26.6)</td>
<td>0.132</td>
</tr>
<tr>
<td>South</td>
<td>4,875 (76.7)</td>
<td>11,116 (66.6)</td>
<td>0.225</td>
</tr>
<tr>
<td>West</td>
<td>128 (2.0)</td>
<td>923 (5.5)</td>
<td>0.185</td>
</tr>
<tr>
<td>Population density, No. (%)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Rural</td>
<td>568 (8.9)</td>
<td>1,551 (9.3)</td>
<td>0.013</td>
</tr>
<tr>
<td>Suburban</td>
<td>1,580 (24.9)</td>
<td>3,970 (23.8)</td>
<td>0.025</td>
</tr>
<tr>
<td>Urban</td>
<td>4,186 (65.9)</td>
<td>11,053 (66.3)</td>
<td>0.008</td>
</tr>
<tr>
<td>Missing</td>
<td>21 (0.3)</td>
<td>105 (0.6)</td>
<td>0.043</td>
</tr>
<tr>
<td>Elixhauser comorbidity index, mean (SD)</td>
<td>1.83 (1.74)</td>
<td>1.64 (1.63)</td>
<td>0.113</td>
</tr>
<tr>
<td>Procedure, No. (%)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Total hip replacement</td>
<td>2,152 (33.9)</td>
<td>5,488 (32.9)</td>
<td>0.020</td>
</tr>
<tr>
<td>Total knee replacement</td>
<td>4,203 (66.1)</td>
<td>11,191 (67.1)</td>
<td>0.020</td>
</tr>
</tbody>
</table>

Abbreviation: SMD, standardized mean difference.

\(^{a}\) We considered SMDs greater than 0.100 to reflect meaningful differences between groups.\(^{19}\) Larger SMDs by geographic region are expected, given the staggered geographic rollout by region.

\(^{b}\) Assessed according to the Centers for Medicare & Medicaid Services beneficiary race and ethnicity code, which reflects data reported to the Social Security Administration.

\(^{c}\) Includes the following races and ethnicities: Asian, Hispanic, North American Native, and other.
$21,964 [95% CI, $21,636-$22,296] vs $22,562 [95% CI, $22,346-$22,779]) (Table 2; eTable 3 in Supplement 1). Time-stratified models for the 3 time periods within a surgical episode (45-day presurgical period, surgical period, and 90-day postsurgical period) are shown in eTable 4 in Supplement 1.

**Postacute Care Use**

In unadjusted analyses, 780 of 6355 bundled episodes (12.3%) were followed by SNF use compared with 4021 of 16,679 (24.1%) for control episodes. In adjusted analyses, participation in the program was associated with a decrease in SNF use (21.3% for bundled episodes vs 25.0% for control episodes; adjusted odds ratio [AOR], 0.81 [95% CI, 0.67-0.98]) compared with control episodes. No statistically significant difference was observed in the use of home health care (42.1% of bundled episodes vs 43.9% for control episodes; AOR, 0.93 [95% CI, 0.72-1.21]) after discharge.

**Surgical Setting**

In unadjusted analyses, outpatient surgical settings were used for 981 of 4334 (22.6%) bundled episodes and 239 of 1763 (13.6%) control episodes after the CMS rule changes. In adjusted analyses,

### Table 2. Adjusted Differences in Total Episodic Spending, Postacute Care Use, Surgical Setting, and Quality, 2012-2019

<table>
<thead>
<tr>
<th>Characteristic</th>
<th>Bundled episodes (95% CI), %</th>
<th>Control episodes (95% CI), %</th>
<th>Adjusted difference (95% CI), %</th>
<th>Adjusted odds ratio (95% CI)</th>
<th>P value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Spending, $</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Total episodic spending</td>
<td>21,964 (21,636 to 22,296)</td>
<td>22,562 (22,346 to 22,779)</td>
<td>−2.7 (−4.1 to −1.3)</td>
<td>NA</td>
<td>&lt;.001</td>
</tr>
<tr>
<td>Postacute care use</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Skilled nursing facility</td>
<td>21.3 (17.9 to 25.1)</td>
<td>25.0 (22.3 to 27.2)</td>
<td>NA</td>
<td>0.81 (0.67 to 0.98)</td>
<td>.03</td>
</tr>
<tr>
<td>Home health care</td>
<td>42.1 (35.4 to 49.2)</td>
<td>43.9 (41.5 to 46.3)</td>
<td>NA</td>
<td>0.93 (0.72 to 1.21)</td>
<td>.60</td>
</tr>
<tr>
<td>Surgical setting</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Outpatient setting use after CMS rule change</td>
<td>14.1</td>
<td>8.4</td>
<td>NA</td>
<td>1.79 (1.53 to 2.09)</td>
<td>&lt;.001</td>
</tr>
<tr>
<td>Quality</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Complications</td>
<td>8.8 (7.3 to 10.5)</td>
<td>8.6 (7.3 to 9.7)</td>
<td>NA</td>
<td>1.02 (0.86 to 1.20)</td>
<td>.84</td>
</tr>
<tr>
<td>Readmissions</td>
<td>4.3 (3.4 to 5.4)</td>
<td>4.6 (3.9 to 4.8)</td>
<td>NA</td>
<td>0.92 (0.75 to 1.13)</td>
<td>.45</td>
</tr>
</tbody>
</table>

Abbreviations: CMS, Centers for Medicare & Medicaid Services; NA, not applicable.

a Adjusted for preindex Elixhauser comorbidity score, age, sex, race and ethnicity, procedure type (total knee replacement vs total hip replacement), time (quarter) fixed effects, and physician practice fixed effects. Total episodic spending included all spending associated with medical and pharmacy claims for that beneficiary starting 45 days before the surgical episode and ending 90 days after it.

b Adjusted for preindex Elixhauser comorbidity score, age, sex, race and ethnicity, and time (year) fixed effects.

c Model included an interaction between bundled payment program participation and the CMS rule change allowing procedures in the outpatient setting in 2018 (for total knee replacement) and 2019 (for total hip replacement). Index-year fixed effects were included instead of quarter-year fixed effects; physician practice fixed effects and clustering were removed.

d Adjusted for preindex Elixhauser comorbidity score, age, sex, race and ethnicity, procedure type (total knee replacement vs total hip replacement), and time (quarter) fixed effects.
bundled episodes were 1.79 times more likely to be performed in the outpatient setting compared with control episodes (14.1% for bundled episodes vs 8.4% for control episodes; AOR, 1.79 [95% CI, 1.53-2.09]) (Table 2).

Quality of Care
In unadjusted analyses, 90-day complications occurred in 525 of 6355 patients (8.3%) for bundled episodes and 1472 of 16 679 patients (8.8%) for control episodes; 90-day readmissions occurred in 258 of 6355 (4.1%) bundled episodes and 784 of 16 679 (4.7%) control episodes. In adjusted analyses, there were no statistically significant differences in quality of care between bundled and control episodes: 90-day complications (8.8% for bundled episodes vs 8.6% for control episodes; AOR, 1.02 [95% CI, 0.86-1.20]) and 90-day readmissions (4.3% for bundled episodes vs 4.6% for control episodes; AOR, 0.92 [95% CI, 0.75-1.13]) (Table 2).

Sensitivity Analyses
Estimated differences in episode spending using Hierarchical Conditions Category scores (~$624 per episode) were similar to those in the primary analysis using Elixhauser scores (~$598 per episode). There was no evidence of an association between bundled payment program participation and changes in beneficiary comorbidity, frailty, disability, or chronic obstructive pulmonary disease status. Full results are provided in eTables 5 to 9 in Supplement 1. Estimated differences in total episodic spending were robust to all sensitivity analyses (eTable 10 in Supplement 1).

Discussion
In this retrospective cross-sectional cohort study, we found that physician practice participation in an LEJR bundled payment program offered by a Medicare Advantage insurer was associated with a decrease in episode spending of 2.7%, equivalent to savings of $598 per episode, without evidence of differences in home health care use or episode quality. Despite a significant increase in Medicare Advantage enrollment during the past decade, to our knowledge, this is the first analysis of an LEJR bundled payment program offered by a Medicare Advantage plan.

The bundled payment program’s reductions in episode spending were associated with reductions in SNF use and an increased use of outpatient surgical settings. There was a strong secular trend toward lower episode spending, reflecting national shifts away from SNF use for patients receiving LEJR (Figure 1; eFigures 4-6 in Supplement 1). Furthermore, because CMS removed total knee replacement surgery from the inpatient-only list on January 1, 2018, some LEJR episodes after that date were performed in outpatient settings for both bundled and control episodes. These findings are consistent with earlier evaluations of bundled payment programs during this time period and with national trends for LEJR episodes, which have documented spending reductions among hospitals not participating in bundled payment programs then.20 Thus, it is notable that the shifts in spending and use associated with the Medicare Advantage insurer’s bundled payment program occurred faster and more steeply.

This evaluation also serves as one of the first analyses of a bundled payment program applied to both inpatient and outpatient surgical settings with conclusions that are useful to incentive design for both private and public insurers.21 Primarily, CMS has offered bundled payments for LEJR procedures completed in the inpatient setting and only recently incorporated site-neutral LEJR into program design. An evaluation of the Comprehensive Care for Joint Replacement program suggested that hospitals participating in the program were less likely to perform surgical procedures in an outpatient setting, in part because outpatient episodes were excluded from the bundle.22 This study offers evidence that site-neutral LEJR can be effectively incorporated into a bundled payment program design to bring about additional program savings without apparently affecting quality of care.
Furthermore, this analysis offers additional evidence on the experience of bundled payment programs for physician practices, which may differ from hospitals or postacute care facilities owing to preexisting relationships with patients, as well as revenue associated with professional fees rather than facility fees. Evaluations of bundled payment programs in traditional Medicare have predominantly focused on the experience of hospitals or postacute care settings. Our findings align with those of 1 earlier study focused on physician practice participants in Bundled Payments for Care Improvement and with analyses of the Medicare Shared Savings Program, which have shown that accountable care organizations led by physician practices experience greater savings. The magnitude of episode spending reductions, reductions in SNF use, and the finding of stable quality in this analysis are all similar to results of studies of bundled payment programs in traditional Medicare. It is notable that the episode spending reductions were similar given that Medicare Advantage plans may use other strategies to select efficient physician practices.

**Limitations**

This study has limitations. First, although we accounted for physician practice selection in the bundled payment program through our study design, because it was an observational study, unobserved confounding is possible. Second, our study was limited to Medicare Advantage beneficiaries enrolled in plans offered by a single insurer, which may limit the generalizability of our findings. Third, because our study design included only physician practices that eventually participated in the bundled payment program, this analysis was designed to measure the association between participation and episode outcomes among those that were "treated" with bundled payments rather than a "mean treatment association"–like estimate of what outcomes would be associated with bundled payment participation among practices that did not participate, which may limit the generalizability of our findings. Fourth, our assessments of quality of care were limited to those measurable in claims data. Fifth, data on reward payments to physician practices were not available, so net savings computations were not feasible. Sixth, participating practices were clustered in certain geographic areas in the southern US.

**Conclusion**

This cross-sectional study found that physician practice participation in an LEJR bundled payment program offered by a Medicare Advantage insurer was associated with a 2.7% decrease in spending without evidence of decrements in quality. This finding suggests that bundled payment programs offered by private payers, including Medicare Advantage plans, may be effective at reducing episode spending while maintaining quality of care.
Concept and design: Sutherland, Boudreau, Bowe, Liao, Flagg, Drzayich Antol, Shrank, Powers, Navathe.

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

Drafting of the manuscript: Boudreau, Huang, Flagg, Cousins, Drzayich Antol, Powers, Navathe.

Critical revision of the manuscript for important intellectual content: Sutherland, Boudreau, Bowe, Liao, Shrank, Powers, Navathe.

Statistical analysis: Sutherland, Bowe, Huang, Powers, Navathe.

Obtained funding: Flagg.

Administrative, technical, or material support: Boudreau, Flagg, Cousins, Shrank, Powers, Navathe.

Supervision: Boudreau, Flagg, Drzayich Antol, Shrank, Powers, Navathe.

Conflict of Interest Disclosures: Dr Sutherland reported employment and equity holdings with Humana. Dr Boudreau reported employment and equity holdings with Humana. Mr Bowe reported employment and equity holdings with Humana. Dr Liao reported receiving honoraria from Comagine Health, Marcus Evans, Brown University, and the Washington Health Alliance outside the submitted work. Mr Flagg reported employment and equity holdings with Humana. Ms Cousins reported receiving grants from Humana during the conduct of the study. Ms Drzayich Antol reported employment and equity holdings with Humana and Andreaessen Horowitz and serving as a director at GetWellNetwork outside the submitted work. Dr Shrank reported employment and equity holdings with Humana and prior employment at Anthem and Fidelity Investments. Dr Navathe reported receiving grants from the Hawaii Medical Service Association, Commonwealth Fund, Robert Wood Johnson Foundation, Donaghue Foundation, Veterans Affairs Administration, Arnold Ventures, United Healthcare, Blue Cross Blue Shield of North Carolina, and Humana; receiving personal fees from Navvis Healthcare, Singapore Ministry of Health, Elsevier Press, Medicare Payment Advisory Commission, Analysis Group, VBID Health, Advocate Physician Partners, Federal Trade Commission, and Catholic Health Services Long Island; having equity in Clarify Health; receiving personal fees from and having board membership in The Scan Group; and having uncompensated board membership in Integrated Services outside the submitted work. No other disclosures were reported.

Funding/Support: This work was supported by a grant to Dr Navathe from Humana.

Role of the Funder/Sponsor: Humana had no role in the decision to submit the manuscript for publication.

Disclaimer: The views expressed in this article represent the authors' views and not necessarily the views or policies of their respective affiliated institutions, the US government, the Department of Veterans Affairs, or the state of Pennsylvania.

Meeting Presentation: This article was presented at the AcademyHealth 2023 Annual Research Meeting; June 25, 2023; Seattle, Washington.

Data Sharing Statement: See Supplement 2.

Additional Contributions: We thank Shane Mains, MBA, Bradley Wright, MBA, and Shannon Burdette, BS, all employees of Humana Inc, for their comments and review of this manuscript. No one received additional compensation beyond usual salary.

REFERENCES


SUPPLEMENT 1.
eAppendix. Supplemental Methods.
eFigure 1. Comparison of Pre-Exposure Spending Outcomes for Early and Late Program Participants, 2012-2015.
eTable 1. Physician Practice Eligibility and Participation by Geography & Year.
eFigure 2. Participant Flow Through the Study.
eTable 2. Unadjusted Spending Results, 2012-2019.
eFigure 3. Distribution of Episodes Over the Study Period, by Exposure Status, 2012-2019
eTable 3. Full Regression Results: Total Spending, 2012-2019.
eTable 4. Episode Time Period Stratified Models
eTable 5. Main Regression Specification for Total Spending With a Beneficiary’s Elixhauser Score Replaced With Their CMS Hierarchical Conditions Category (CMS-HCC) Score
eTable 6. Main Regression Specification With a Beneficiary’s Elixhauser Score as the Dependent Variable
eTable 7. Main Regression Specification With a Beneficiary’s Frailty Index as the Dependent Variable
eTable 8. Main Logistic Regression Specification With a Beneficiary’s Disability Status as the Dependent Variable
eTable 9. Main Logistic Regression Specification With a Beneficiary’s COPD Status as the Dependent Variable
eTable 10. Total Episodic Spending Model Sensitivity Analyses
eFigure 4. Change in SNF Utilization for Control Episodes Over Time (2012-2019)
eFigure 5. Change in Post 90-Day Time Period Spending for Control Episodes Over Time (2012-2019)
eFigure 6. Change in Surgical Time Period Spending for Control Episodes Over Time (2012-2019)

SUPPLEMENT 2.
Data Sharing Statement