Arthroplasty Today 5 (2019) 119–125

Contents lists available at ScienceDirect

Arthroplasty Today

journal homepage: http://www.arthroplastytoday.org/

Original research

Short-term functional recovery after total joint arthroplasty is unaffected by bundled payment participation

Joshua K. Johnson, DPT [a,*], Jill A. Erickson, PA-C [b], Caitlin J. Miller, DPT [a], Julie M. Fritz, PT, PhD [a], Robin L. Marcus, PT, PhD [a], Christopher E. Pelt, MD [b]

[a] *Department of Physical Therapy and Athletic Training, University of Utah, Salt Lake City, UT, USA*
[b] *Department of Orthopaedics, University of Utah, Salt Lake City, UT, USA*

ARTICLE INFO

Article history:
Received 22 July 2018
Received in revised form
22 December 2018
Accepted 29 December 2018
Available online 7 February 2019

Keywords:
Value
Alternative payment models
Physical function
Patient outcomes
Arthroplasty

ABSTRACT

Background: Bundled payment models for lower extremity total joint arthroplasty (TJA) aim to improve value by decreasing costs via efficient care pathways. It is unclear how such models affect patient-centered outcomes such as functional recovery. We aimed to determine whether participation in bundled payment for TJA negatively affects patients' functional recovery.
Methods: All patients, regardless of payer, undergoing elective TJA between July 2014 and December 2016 were identified retrospectively and categorized into prebundle (n = 680) and postbundle (n = 1216) cohorts. Mixed-effects linear regression and Wald postests were used to test for differences in patients' functional recovery during the hospital period and over 12 months after TJA between cohorts. We also used multivariate regression to test for differences in hospital length of stay (LOS) and postacute care (PAC) facility use between cohorts.
Results: Compared with the prebundle cohort, patients in the postbundle cohort demonstrated a small and nonmeaningful difference in the trajectory of functional recovery in the hospital [$\chi^2(3) = 31.3$, $P < .01$] and no difference in the 12 months after TJA [$\chi^2(3) = 3.9$, $P = .28$]. They had a 0.4-day shorter hospital LOS (95% confidence interval: −0.5, −0.3) and decreased odds for PAC facility use (adjusted odds ratio = 0.3; 95% confidence interval: 0.2, 0.4).
Conclusions: Participation in bundled payment for TJA was not associated with significant changes in patients' functional recovery, an important patient-centered outcome. For the postbundle cohort, hospital LOS and PAC facility use were decreased, consistent with previous studies describing cost-saving strategies in bundled payment. These findings support the need for an ongoing study of the long-term sustainability of these value-based payment models.

Introduction

Current health-care reform in the United States aims to shift away from fee-for-service payment models toward reimbursement of value-based care [1,2]. Greater value in health-care delivery is achieved when the costs of care are minimized and patient outcomes are maximized [2,3]. The Centers for Medicare and Medicaid Services introduced bundled payment programs as part of the Affordable Care Act with the intent of improving value by decreasing costs via fixed pricing for an episode of care [4]. This incentivizes hospitals and care providers to improve efficiency in care processes. Elective lower extremity total joint arthroplasty (TJA) of the hip (THA) or knee (TKA) is known to be amenable to established care processes [5] and is the most common inpatient surgery underwent by Medicare beneficiaries [6] and so was selected for the trial of bundled payment models beginning in 2009.

Studies have reported successful cost-containment thus far from bundling payment for TJA. Specifically, costs have been reduced

One or more of the authors of this article have disclosed potential or pertinent conflicts of interest, which may include receipt of payment, either direct or indirect, institutional support, or association with an entity in the biomedical field which may be perceived to have potential conflict of interest with this work. For full disclosure statements refer to https://doi.org/10.1016/j.artd.2018.12.003.

* Corresponding author. Department of Physical Therapy and Athletic Training, University of Utah, 520 Wakara Way, Salt Lake City, UT 84108, USA. Tel.: +1 216 903 0621.

E-mail address: joshua.johnson@hsc.utah.edu

https://doi.org/10.1016/j.artd.2018.12.003
2352-3441/

using strategies such as standardization of surgical implants and processes [7], reduction of postoperative hospital length of stay (LOS) [7-13], and decreased use of postacute care (PAC) services [7,9,14,15] in skilled nursing and inpatient rehabilitation facilities. Importantly, it has also been shown that there is no change in unplanned readmission, emergency department use, or mortality in the presence of bundled payment for TJA [9]. These are important outcomes being monitored by the Centers for Medicare and Medicaid Services.

However, given that TJA is an elective procedure from which patients are seeking a reduction of pain and improvement in function, these outcomes alone are insufficient to examine value from a patient-centered perspective [16-18]. Outcomes of physical function in association with bundled payment for TJA have only been reported in cross-sectional evaluations [9]. Further research is needed evaluating longitudinal trajectories of patient-centered outcomes after TJA.

In an effort to maximize value for patients undergoing TJA, our health-care system opted into the Bundled Payment Care Improvement (BPCI) Model 2 in July 2015. Although the BPCI aims at increasing value specifically for a subset of Medicare beneficiaries, we enhanced care pathways for all patients undergoing TJA even before our BPCI enrollment. These efforts, described in full elsewhere [13,15,19], resulted in decreased hospital LOS, decreased PAC facility use, and reductions in 30-day and 90-day readmission or reoperation, consistent with findings from other organizations [7-12]. These are important outcomes to the value equation as they indicate improved—or maintained—patient outcomes and likely decreases in costs. To further examine potential BPCI-associated value from a patient-centered perspective, the primary aim of the present study was to test whether the trajectory of post-TJA physical function was negatively affected for patients undergoing TJA after we began participating in the BPCI. To understand the extent that BPCI participation potentially influenced cost changes via modified patient care pathways, we also report on hospital LOS and PAC use for the specific cohort of patients included in this study.

Material and methods

Care pathway improvements associated with BPCI enrollment

In preparation for our enrollment in the BPCI on July 1, 2015, several improvements were made to the care pathway for all patients undergoing TJA, regardless of BPCI eligibility. Among these were an increased emphasis on earlier postoperative ambulation beginning in 2013 that involved the addition of a swing shift (11 AM-8 PM) for inpatient orthopedic physical therapy (PT) staff [13]. In preparation for bundled payment participation, refinements to this staffing model aimed to facilitate an appropriate reduction in post-TJA hospital LOS. In addition, stepwise implementation of a comprehensive patient education and management program began in June 2015 and culminated in November 2015 [15]. The program included preoperative education in a formal classroom setting, postoperative follow-up via telelphone calls from a care navigator, and a 24-hour provider on-call system, all aimed at providing patients with sufficient resources to limit adverse events without requiring care in a PAC facility. We also implemented processes for standardized clinical measurement of physical function in 2014. Standard measurement across outpatient clinic settings occurs before and after the operation using the computer adaptive test of the patient-reported outcome measurement information system physical function scale (PF CAT). Immediate postoperative function is assessed in the hospital using the Activity Measure for Post-acute Care “6-clicks” basic mobility short form (AM-PAC-Mobility).

Data

We collected patient data retrospectively from our institution’s enterprise data warehouse (EDW), a tool that combines administrative data with clinical data from provider notes recorded in the electronic medical record. This includes AM-PAC-Mobility and PF CAT assessment scores, both of which are recorded directly into the EMR at the time of assessment. The initial data extraction included data of all patients meeting the study’s inclusion criteria, regardless of their BPCI eligibility.

Study population

Episodes of all primary TJA procedures initially identified in the EDW were of patients whose surgery date was between July 1, 2014, and December 31, 2016, at our academic medical center’s primary hospital. Unique surgical episodes, rather than unique patients, were considered for the analysis. Surgical cases were identified using Medicare severity diagnosis-related group (MS-DRG) codes 469 and 470. Surgeries coded as “Urgent” or “Trauma Urgent” were excluded so that only elective procedures were included for analysis. Patients were also excluded if they did not have at least one functional mobility score recorded after surgery, using either the AM-PAC-Mobility or the PF CAT. Remaining patient episodes were divided into 2 cohorts based on the date of their TJA procedure. The prebundle cohort included those patients who underwent elective TJA between July 1, 2014, and June 30, 2015, and the postbundle cohort included those patients who underwent elective TJA between July 1, 2015, and December 31, 2016. The prebundle cohort was limited to 12 months as we did not implement the use of the AM-PAC-Mobility until July 2014.

Predictor variables

Cohort inclusion (prebundle or postbundle period) was the primary predictor variable for all analyses. The contribution of various patient characteristics to the outcomes of interest was also considered. These included sex, age, body mass index (BMI), the presence of Medicare as the primary payer, comorbidity burden measured via the Charlson Comorbidity Index (CCI) [20], diagnostic severity as indicated by the Medicare Severity Diagnosis Related Group weight, and preoperative physical function measured using the PF CAT. Consistent with other studies examining the outcomes associated with bundled payment in TJA [7,9,10,21], patients with TKA and THA were analyzed together, but the specific procedure was also considered as a variable input in statistical analyses.

Patient-centered functional recovery outcomes

For the primary study objective, we retrospectively identified and analyzed individual patients’ trajectory of functional recovery based on general reports of function as rated by the patient or a clinical proxy. This was completed separately for the postoperative hospital period and in the 12 months after TJA. The functional status in the hospital was assessed using the AM-PAC-Mobility. This is a clinician-rated measure of a patient’s ability to complete basic mobility tasks, validated for use across the spectrum of hospitalized patients [22]. Scores on the T-scale of the AM-PAC instrument range from 0 to 100, with higher scores indicating greater independence with functional mobility. Timing of AM-PAC-Mobility assessment was categorized into 4 distinct time points relative to the time of surgery: Day 0 (assessment recorded between 0 and 12 hours after TJA), Day 1 (12-24 hours), Day 2 (24-48 hours), and Day 3+, which captured the final overall score recorded in the hospital for patients with an LOS of 48 hours or more.

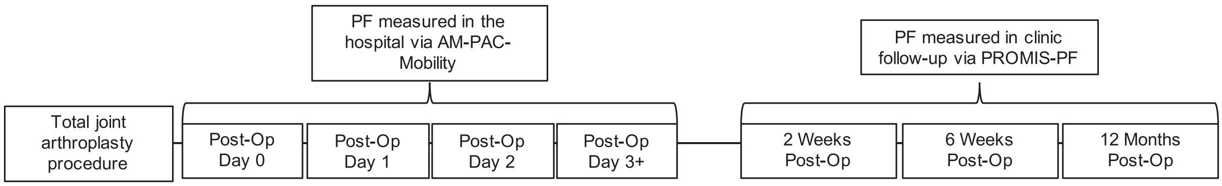

Figure 1. Time points of physical function assessment by the measure used.

The trajectory of physical function in the 12 months after TJA was assessed for each patient using the PF CAT. This self-report of physical function has been validated for assessment of patients seeking care in ambulatory care settings [23] and specifically those with orthopedic conditions [24]. The PF CAT is reported as a T-score ranging from 0 to 100, with a population mean of 50 and a standard deviation of 10. Higher scores indicate greater physical function. The study time points were associated with the standard timing of clinic visits for patients undergoing TJA at our institution: preoperative, 2 weeks after operation, 6 weeks after operation, and 12 months after operation. Figure 1 displays the time points of assessment with either PF instrument.

The preoperative score used for each patient was that score recorded closest to, but before, the date of surgery. Postoperative 2-week scores included any score within 5 and 27 days after TJA, with the score used for analysis being that which was closest to 14 days after surgery. Scores at the 6-week time point included any score recorded within 30 and 90 days after TJA; the score recorded closest to 42 days was the one included in analysis. The 12-month time point captured a patient's final score, recorded at any time between 270 and 450 days after TJA.

Care utilization outcomes

In the absence of cost data, we tested for differences in the extent to which clinical practice patterns differed between the prebundle and postbundle periods using variables related to care utilization. These included hospital LOS for the perioperative period and discharge disposition location. Patients were categorized as being discharged to a PAC facility if discharged to either a skilled nursing or inpatient rehabilitation facility.

Statistical analysis

Potential differences in patient characteristics between groups were assessed using an independent *t*-test for continuous variables and a chi-square test for categorical variables. Patient-centered outcome and care-utilization variables were evaluated using the regression method appropriate for the level of measurement of the outcome variable. Each of the patient characteristic variables (Table 1) was initially included as confounders in each regression analysis. The final model for each analysis was derived using backward variable selection with a conservative significance threshold (0.20) for variable inclusion [25]. All analyses were conducted in Stata version 15.1 (StataCorp, College Station, TX).

Measures of a patient's physical function recovery, both in the hospital and over 12-month follow-up, were analyzed separately using mixed-effects linear regression. This is a multilevel model approach that can fit individual trajectories of function for each patient while controlling for both fixed and random effects at specified levels in longitudinal data [26]. For both in-hospital and 12-month follow-up, all patient characteristics described in Table 1 were initially included as fixed effects in the model. The TJA type (THA or TKA) and the timing of scores relative to surgery were included as random effects and nested within individual patients who were within the bundle cohort variable. For both the AM-PAC-Mobility and the PF CAT models, we included an interaction term of bundle cohort × time point of interest to determine the significance of the difference in trajectories between the patients in the 2 cohorts. From regression results, we used marginal estimation to determine an adjusted mean and 95% confidence interval (CI) for functional scores at each time point and a post-estimation Wald test to assess for differences in the trajectory of functional recovery between patients in each cohort.

Mixed-effects linear regression is especially useful when testing the relationship between an outcome and an individual rather than making inferences about the population [27], as is the case in this study. It is also compatible with maximum-likelihood estimation (MLE), which we used to estimate missing values for variables in the model, particularly missing functional scores, to limit the selection bias in the sample. MLE has been shown to outperform methods of data imputation for longitudinal data analysis, including outcome data [28,29]. We also completed sensitivity analyses for each functional outcome in which we included only those patients who had scores recorded for at least 3 of the AM-PAC-Mobility or the PF CAT time points, respectively, in an effort to minimize the dependence of the results on MLE.

Table 1
Patient characteristics by cohort.

Variable	Prebundle	Postbundle	*P* value
Sample size, n	680	1216	-
TKA, n (%)	416 (61.2%)	682 (56.1%)	.03
Male, n (%)	294 (43.2%)	516 (42.4%)	.74
Age, mean (SD)	62.4 (11.4)	62.8 (11.7)	.44
BMI, mean (SD)	30.9 (6.8)	30.4 (6.6)	.10
CCI, mean (SD)	1.7 (2.3)	1.8 (2.3)	.46
MS-DRG weight, mean (SD)	2.2 (0.2)	2.1 (0.2)	<.01
Preoperative PF CAT score, mean (SD)	36.4 (6.4)	36.9 (6.6)	.20
Medicare primary, n (%)	293 (43.1%)	548 (45.1%)	.41

n, sample size; SD, standard deviation.

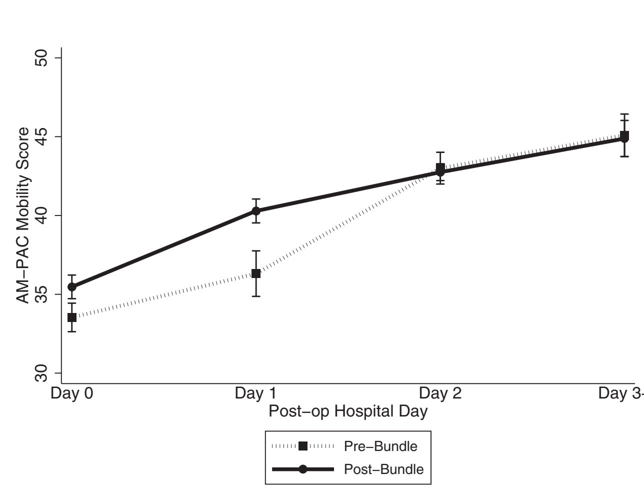

Figure 2. Adjusted trajectories of physical function in the hospital immediately after total joint arthroplasty (with 95% confidence intervals).

Table 2
Functional outcome scores at each time point by cohort.

Assessment used	Time point	Observed score, mean (95% CI)		Adjusted score, mean (95% CI)		Patients with recorded score, n (%)	
		Prebundle	Postbundle	Prebundle	Postbundle	Prebundle	Postbundle
AM-PAC-Mobility	Day 0	32.1 (31.4, 32.7)	35.4 (35.0, 35.8)	33.5 (32.6, 34.4)	35.5 (34.7, 36.2)	459 (67.5%)	990 (81.4%)
	Day 1	35.1 (33.1, 37.1)	40.8 (39.9, 41.8)	36.3 (34.9, 37.8)	40.3 (39.5, 41.0)	101 (14.9%)	378 (31.1%)
	Day 2	43.0 (41.6, 44.4)	43.2 (42.6, 43.8)	43.0 (42.0, 44.0)	42.7 (42.2, 43.3)	207 (30.4%)	778 (64.0%)
	Day 3+	44.0 (43.1, 45.0)	44.2 (43.5, 44.9)	45.1 (43.7, 46.4)	44.9 (43.7, 46.0)	242 (35.6%)	468 (38.4%)
PF CAT	Preoperative	36.4 (35.9, 37.0)	36.8 (36.4, 37.2)	36.4 (35.8, 37.0)	36.2 (35.6, 36.8)	619 (91.0%)	1106 (91.0%)
	2 weeks	32.2 (31.6, 32.9)	33.1 (32.6, 33.6)	32.4 (31.9, 32.9)	32.5 (32.0, 32.9)	491 (72.2%)	742 (61.0%)
	6 weeks	39.7 (39.1, 40.3)	40.1 (39.6, 40.5)	40.2 (39.7, 40.6)	39.6 (39.2, 40.0)	585 (86.0%)	990 (81.4%)
	12 months	43.6 (42.8, 44.5)	43.1 (42.4, 43.8)	45.3 (43.8, 46.7)	44.4 (43.0, 45.7)	337 (49.6%)	550 (45.2%)

CI, confidence interval; n, sample size.

The difference in postoperative hospital LOS was tested using generalized gamma regression, which is a generalized linear model with a log link and gamma family and is robust to skewed outcome data such as hospital LOS [30]. Discharge to a PAC facility was dichotomized as a yes or no variable and was analyzed using multiple logistic regression to estimate an adjusted odds ratio. This ratio indicates the odds of an outcome event occurring in the postbundle cohort compared with the prebundle cohort after adjusting for confounding effects.

Results

Study population characteristics

There were 1947 unique surgical episodes initially identified from the EDW. Of these, 44 were coded as urgent, 6 were missing functional scores at all time points, and one had undergone a second procedure within the same hospital period. These 51 cases were excluded, leaving 1896 TJA episodes for 1666 unique patients available for analysis, 680 episodes in the prebundle cohort, and 1216 episodes in the postbundle cohort. These included 1098 (57.9%) episodes of TKA. The overall mean (standard deviation) age was 62.6 (11.6) years, and most patients were women (57.3%). There was a statistical difference observed between the prebundle and postbundle cohorts for the proportion of patients undergoing TKA vs THA and for MS-DRG weight (Table 1). These differences, and other patient characteristics determined to be statistically associated with each outcome, were accounted for in our adjusted models.

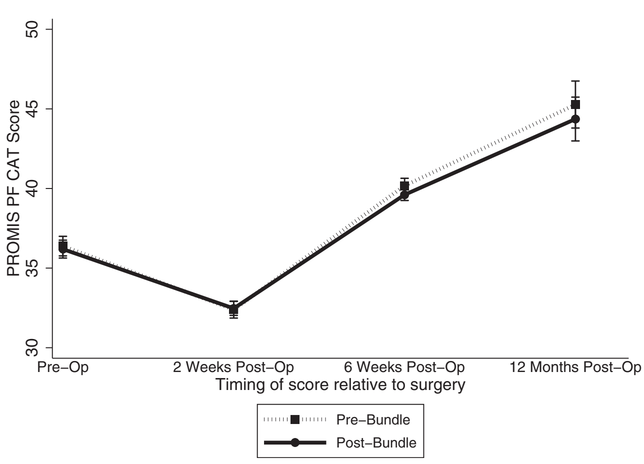

Figure 3. Adjusted trajectories of physical function over 12-month follow-up (with 95% confidence intervals).

Patient-centered functional recovery

As shown in Figure 2, during their postoperative hospitalization period, patients in the postbundle cohort demonstrated a statistical difference in functional recovery trajectory compared with those in the prebundle cohort [$\chi^2(3) = 31.3$, $P < .01$]. The estimated means of AM-PAC-Mobility scores indicate higher levels of independent physical function among those in the postbundle cohort than in the prebundle cohort at day 0 and 1, but not at day 2 or 3 (Table 2). The greatest difference between the 2 cohorts was observed at day 1 where the estimated mean AM-PAC-Mobility score for the prebundle cohort was 36.3 (95% CI: 34.9, 37.8) and that for the postbundle cohort was 40.3 (95% CI: 39.5, 41.0), with an adjusted mean difference of 4.0 points.

Longitudinal measures of the PF CAT indicate that there was no difference in the trajectory of functional recovery over 12 months of post-TJA follow-up [$\chi^2(3) = 3.9$, $P = .28$; Fig. 3]. No difference was observed between the cohorts at any individual time point.

Missing scores and sensitivity analyses

Considering all time points and the full sample size, 7584 scores from the AM-PAC-Mobility and PF CAT should have been recorded. However, 3810 (50.2%) AM-PAC-Mobility scores and 2164 (28.5%) PF CAT scores were missing and therefore estimated using MLE. Table 2 summarizes the proportion of scores collected for each measure at each time point.

There were 457 patients (68 in the prebundle cohort and 389 in the postbundle cohort), with at least 3 AM-PAC-Mobility scores recorded. As shown in Table 3, there was no difference between cohorts in any measured characteristic for patients included in the sensitivity analysis. As with the full sample, there was a difference between prebundle and postbundle trajectories of functional recovery in the hospital period [$\chi^2(3) = 60.8$; $P < .01$; Fig. 4]. The largest difference in AM-PAC-Mobility scores for patients included

Table 3
Sensitivity analyses results: patient characteristics by cohort for patients with at least 3 AM-PAC-Mobility scores recorded.

Variable	Prebundle	Postbundle	*P* value
Sample size, n	68	389	-
TKA, n (%)	38 (55.9%)	243 (62.5%)	.30
Male, n (%)	24 (35.3%)	127 (32.6%)	.67
Age, mean (SD)	63.2 (13.0)	65.0 (11.7)	.27
BMI, mean (SD)	31.0 (6.4)	31.0 (6.9)	.94
CCI, mean (SD)	1.4 (2.2)	2.1 (2.5)	.05
MS-DRG weight, mean (SD)	2.2 (0.2)	2.1 (0.3)	.29
Preoperative PF CAT score, mean (SD)	34.8 (5.9)	35.8 (6.6)	.30
Medicare primary, n (%)	30 (44.1%)	219 (56.3%)	.06

n, sample size; SD, standard deviation.

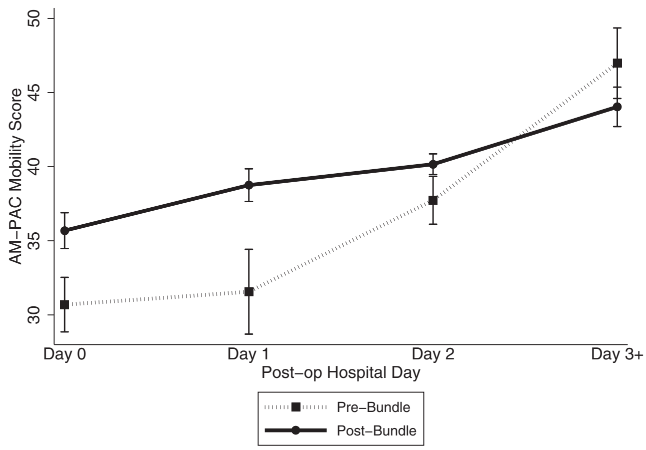

Figure 4. Sensitivity analysis results. Adjusted trajectories of physical function in the hospital immediately after total joint arthroplasty (with 95% confidence intervals) for those patients with at least 3 AM-PAC-Mobility scores recorded.

in the sensitivity analysis was observed at day 1 after TJA. At that point, the mean score for the prebundle cohort was 31.6 (95% CI: 28.7, 34.4) and that for the postbundle cohort was 38.8 (95% CI: 37.7, 39.9; Table 4).

The prebundle and postbundle cohorts included 287 and 450 patients, respectively, that had at least 3 PF CAT scores recorded and were therefore included in the sensitivity analysis. These patients differed statistically only for MS-DRG weight (Table 5), which was held constant along with other confounders in the mixed-effects linear model for this sensitivity analysis. For this smaller patient sample, there remained no difference between cohorts in the trajectory of functional recovery [$\chi^2(3) = 4.2$; $P = .24$] as shown in Figure 5.

Care utilization

The median (interquartile range) perioperative hospital LOS for the prebundle cohort was 2.3 (2.0, 3.1) days, and this decreased for the postbundle cohort to 2.1 (1.3, 2.4) days. The unadjusted gamma regression indicated a 0.5 (95% CI: −0.6, −0.4)-day reduction in hospital LOS for the postbundle cohort. After controlling for gender, age, BMI, the CCI, MS-DRG weight, preoperative PF CAT score, and whether Medicare was the primary payer, there was a 0.4 (95% CI: −0.5, −0.3)-day reduction in LOS.

There was also decreased odds for discharge to a PAC facility for patients in the postbundle cohort (adjusted odds ratio = 0.3; 95% CI: 0.2, 0.4), with 10.9% being discharged to a PAC facility relative to 26.9% in the prebundle cohort. All other patients in the study

Table 5
Sensitivity analyses results: patient characteristics by cohort for patients with at least 3 PF CAT scores recorded.

Variable	Prebundle	Postbundle	*P* value
Sample size, n	287	450	-
TKA, n (%)	176 (61.3%)	257 (57.1%)	.26
Male, n (%)	129 (44.9%)	175 (38.9%)	.50
Age, mean (SD)	61.2 (11.6)	62.0 (11.1)	.30
BMI, mean (SD)	30.7 (6.2)	30.4 (6.2)	.45
CCI, mean (SD)	1.6 (2.3)	1.8 (2.2)	.39
MS-DRG weight, mean (SD)	2.2 (0.2)	2.1 (0.2)	.01
Preoperative PF CAT score, mean (SD)	36.4 (7.0)	36.7 (6.4)	.58
Medicare primary, n (%)	110 (38.3%)	192 (42.7%)	.24

n, sample size; SD, standard deviation.

sample were discharged home, either with or without home health services.

Discussion

Regarding the need to evaluate patient-centered outcomes as a consideration of value-based care, we sought in this study to test whether participation in a bundled payment program had an adverse effect on the trajectory of physical function recovery for all patients undergoing elective TJA. The analysis shows that patients' recovery of physical function was statistically—but not clinically—different between the prebundle and postbundle cohorts during the in-hospital period, but not in the 12 months after TJA. This is the case despite the fact that hospital LOS was decreased and odds for discharge to a PAC facility were decreased for the postbundle cohort.

Although patients in the postbundle cohort demonstrated higher levels of physical function than those in the prebundle cohort in the immediate days after surgery, this difference does not exceed the minimal detectable change for the AM-PAC-Mobility of 4.7 points [22]. The greatest difference observed between the 2 cohorts was only 4.0 points (at hospital day 1), which may simply be a result of a measurement error. It is, therefore, not likely to be clinically meaningful, although a minimal clinically important difference for the AM-PAC-Mobility has not been estimated. Regardless, recovery trajectories converge and are nearly identical between cohorts for days 2 and 3 after TJA. This similarity in functional recovery continues through the 12 months after TJA.

Hospital LOS and the use of PAC facilities were both decreased in the postbundle cohort compared with the prebundle cohort, outcomes that are consistent with the results of previous studies at our institution [13,15,19]. These are strategies that have been shown in previous studies to be associated with cost savings. Dummit and colleagues [9] determined from Medicare claims data that there is a cost decrease of $1166 per TJA episode associated with bundled

Table 4
Sensitivity analyses results: functional outcome scores at each time point by cohort for patients with at least 3 AM-PAC-Mobility or PF CAT scores recorded, respectively.

Assessment used	Time point	Observed score, mean (95% CI)		Adjusted score, mean (95% CI)		Patients with recorded score, n (%)	
		Prebundle	Postbundle	Prebundle	Postbundle	Prebundle	Postbundle
AM-PAC-mobility	Day 0	29.0 (27.3, 30.8)	34.5 (33.9, 35.2)	30.7 (28.9, 32.5)	35.7 (34.5, 36.9)	60 (88.2%)	364 (93.6%)
	Day 1	29.5 (26.6, 32.4)	38.2 (37.0, 39.5)	31.6 (28.7, 34.4)	38.8 (37.7, 39.9)	15 (22.1%)	134 (34.4%)
	Day 2	37.3 (35.9, 38.7)	40.1 (39.3, 40.9)	37.7 (36.1, 39.3)	40.2 (39.5, 40.9)	66 (97.1%)	336 (86.4%)
	Day 3+	47.5 (45.2, 49.7)	44.9 (44.0, 45.7)	47.0 (44.6, 49.4)	44.0 (42.7, 45.4)	64 (94.1%)	343 (88.2%)
PF CAT	Preoperative	36.4 (35.6, 37.2)	36.7 (36.0, 37.3)	36.3 (35.4, 37.3)	36.2 (35.3, 37.1)	271 (94.4%)	436 (96.9%)
	2 weeks	32.1 (31.1, 33.2)	32.9 (32.1, 33.7)	32.3 (31.5, 33.1)	32.8 (32.1, 33.4)	232 (80.9%)	329 (73.1%)
	6 weeks	39.7 (38.8, 40.5)	39.9 (39.2, 40.6)	40.1 (39.4, 40.8)	39.7 (39.1, 40.2)	270 (94.1%)	424 (94.2%)
	12 months	44.2 (42.5, 45.9)	42.6 (41.3, 44.0)	45.1 (42.2, 48.1)	43.5 (40.8, 46.3)	88 (30.7%)	161 (35.8%)

CI, confidence interval; n, sample size.

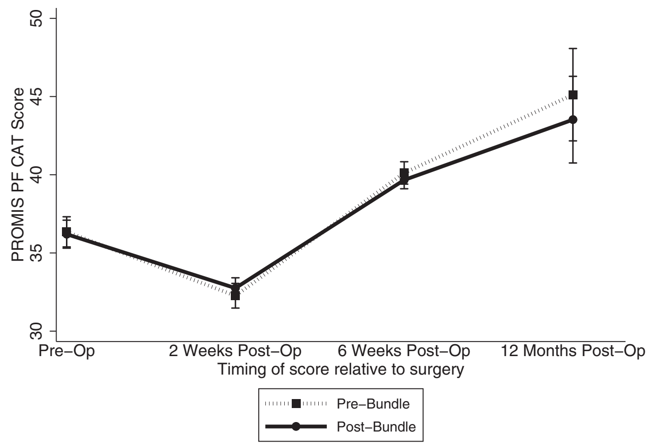

Figure 5. Sensitivity analysis results. Adjusted trajectories of physical function over 12-month follow-up (with 95% confidence intervals) for those patients with at least 3 PF CAT scores recorded.

payment. Likewise, in their analysis of bundled payment participation in a single health system, Navathe et al [7] showed a $5577 decrease in hospital costs per TJA episode. Both studies cite decreased use of PAC as a primary driver of cost savings. Similarly, decreased hospital LOS has been shown to contribute significantly to cost reductions for patients after TJA [8,31,32].

Although these novel findings are encouraging, other unintended consequences of bundled payment for TJA should be considered. First, decreased hospital LOS and the use of PAC facilities create more rapid pathways for patients to be at home after TJA. It is unclear whether these modified care pathways contribute to a disproportionatel shift of care burdens from health-care providers to patients and caregivers at home. Similarly, it may be that patients and/or their caregivers incur increased costs associated with a potential shift in care responsibilities. Whether these consequences exist is not clear and warrant exploration in future studies. Second, the extent to which bundled payments influence decisions regarding the appropriate selection of surgical candidates has not been examined. Without considering the specific effects of bundled payment models, overutilization of TJA has been recognized [33,34]. The incentives associated with bundled payment models could shift the utilization of TJA toward healthier patients and away from patients with greater risk for complications. It remains to be seen whether or not this has occurred or will occur on a wide scale. For our analysis, the nonsignificant differences in the baseline age, BMI, CCI, and preoperative PF CAT scores between the prebundle and postbundle cohorts suggest that no such shift in patient selection has occurred at our institution. However, understanding the potential for these unintended consequences broadly requires additional studies.

Limitations

Our study has several limitations to consider. First, this was a retrospective observational study, thereby limiting any potential for causal inference from the findings. We explain only observed associations between participation in a bundled payment model and the outcomes of interest. The retrospective nature of the study, using our own system's EDW as the sole data source, limits our ability to capture events or outcomes that could further indicate a patient's success in regaining function after TJA. Furthermore, the EDW, like other data repositories [35], may contain data errors that are not easily identifiable at the stage of research investigation. Thus, there is some assumption regarding the accuracy of the data being used.

Second, the implementation of revised care pathways was fluid in its timing and did not correlate precisely with the timing of our BPCI enrollment. For example, the processes associated with earlier post-TJA out-of-bed activity were implemented during the prebundle period. The other noted pathway changes were initially implemented in the final months of the prebundle period and finalized in the early months of the postbundle period. Thus, there may be some spillover effects among patients in either the prebundle or postbundle cohort. However, it is clear from the observed differences between prebundle and postbundle cohorts in hospital LOS and discharge disposition that the expected changes to patient care pathways were not well established until after BPCI enrollment.

Third, the AM-PAC-Mobility has been validated in a heterogenous population of hospitalized patients [22], but not specifically for patients undergoing TJA. There may be limitations in the sensitivity of PF assessment using the AM-PAC-Mobility for this particular population that are not yet understood.

Fourth, the external validity of the findings is limited due to the fact that data were used for patients from only one academic health system. There may be patient or practice characteristics at our institution that differ from those of other health systems, including the timing of postoperative clinical follow-up. Similarly, we used wide date ranges around the standard follow-up time points at our institution to capture a greater number of assessments, but doing so may capture erroneous scores. To limit the possibility of these errors, we identified and retained only the scores recorded for each patient closest to the time point of interest. There may have also been erroneous scores captured for the 230 TJA episodes that were for nonunique patients in the data set. It may be that PF CAT scores captured in the EDW did not necessarily correlate with the surgical episode with which it was paired in the analysis.

Fifth, the significant volume of missing AM-PAC-Mobility and PF CAT scores limits the reliability of the findings. Although we used statistical modeling appropriate for estimating missing outcome data and conducted sensitivity analyses for patient samples with higher proportions of complete data, estimation errors could still be present.

Finally, other variables for which we did not account could influence the observed estimates. For example, we did not account for the use of PT interventions at any point during the course of care. As PT is commonly used to manage functional deficits associated with TJA, its influence should be considered more fully. Work in our own system shows that PT treatment provided on the day of surgery has contributed to modest decreases in LOS and hospital costs. [13] In addition to PT treatment in the hospital, the use of PT in home health or outpatient settings also deserves examination, from the perspective of both cost and patient outcomes.

Conclusions

Value in health care is achieved when patient-centered outcomes are maintained or improved in the presence of cost-saving care. This analysis provides a promising view of the value associated with bundled payments for TJA. Results suggest that the trajectory of patients' physical function recovery is not adversely affected in the presence of care pathways known to reduce costs. Although encouraging, the full scope of these pathways should continue to be analyzed to maximize value-based care delivery by identifying the interventions that can simultaneously reduce costs and improve patient outcomes. In addition, the total impact of bundled payment models for TJA on patients and their caregivers should be more fully understood. Continued effort toward this understanding can ensure that health-care systems are providing optimal benefit for patients with the greatest effectiveness and efficiency.

Acknowledgments

The authors express appreciation to Mike Anderson and Greg Stoddard for their insights pertaining to the statistical analyses associated with this work.

Sources of funding: This investigation was supported, in part, by an American Physical Therapy Association, United States Innovations 2.0 research grant and Florence P. Kendall and Promotion of Doctoral Study scholarships from the Foundation for Physical Therapy, United States. The funding agencies did not play any role in the conduct or report of this study.

References

[1] Miller HD. From volume to value: better ways to pay for health care. Health Aff (Millwood) 2009;28(5):1418.
[2] Porter ME. Value-based health care delivery. Ann Surg 2008;248(4):503.
[3] Porter ME. What is value in health care? N Engl J Med 2010;363(26):2477.
[4] Press MJ, Rajkumar R, Conway PH. Medicare's new bundled payments: design, strategy, and evolution. JAMA 2016;315(2):131.
[5] Lansky D, Nwachukwu BU, Bozic KJ. Using financial incentives to improve value in orthopaedics. Clin Orthop Relat Res 2012;470(4):1027.
[6] Comprehensive care for joint replacement model. https://innovation.cms.gov/initiatives/cjr. Accessed April 17, 2017.
[7] Navathe AS, Troxel AB, Liao JM, et al. Cost of joint replacement using bundled payment models. JAMA Intern Med 2017;177(2):214.
[8] Clair AJ, Evangelista PJ, Lajam CM, Slover JD, Bosco JA, Iorio R. Cost analysis of total joint arthroplasty readmissions in a bundled payment care improvement initiative. J Arthroplasty 2016;31(9):1862.
[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(12):1267.
[10] Dundon JM, Bosco J, Slover J, Yu S, Sayeed Y, Iorio R. Improvement in total joint replacement quality metrics: year one versus year three of the bundled payments for care improvement Initiative. J Bone Joint Surg Am 2016;98(23):1949.
[11] McLawhorn AS, Buller LT. Bundled payments in total joint replacement: keeping our care affordable and high in quality. Curr Rev Musculoskelet Med 2017;10(3):370.
[12] Kiskaddon EM, Lee JH, Meeks BD, Barnhill SW, Froehle AW, Krishnamurthy A. Hospital discharge within 1 day after total joint arthroplasty from a veterans affairs hospital does not increase complication and readmission rates. J Arthroplasty 2017;33(5):1337.
[13] Pelt CE, Anderson MB, Pendleton R, Foulks M, Peters CL, Gililland JM. Improving value in primary total joint arthroplasty care pathways: changes in inpatient physical therapy staffing. Arthroplast Today 2017;3(1):45.
[14] Mouille B, Higuera C, Woicehovich L, Deadwiler M. How to succeed in bundled payments for total joint replacement. NEJM Catalyst. 2016. https://catalyst.nejm.org/how-to-succeed-in-bundled-payments-for-total-joint-replacement/. Accessed December 2, 2017.
[15] Pelt CE, Gililland JM, Erickson JA, Trimble DE, Anderson MB, Peters CL. Improving value in total joint arthroplasty: a comprehensive patient education and management program decreases discharge to post-acute care facilities and post-operative complications. J Arthroplasty 2018;33(1):14.
[16] Porter ME, Larsson S, Lee TH. Standardizing patient outcomes measurement. N Engl J Med 2016;374(6):504.
[17] de Achaval S, Kallen MA, Amick B, et al. Patients' expectations about total knee arthroplasty outcomes. Health Expect 2016;19(2):299.
[18] Franklin PD, Li WJ, Ayers DC. The Chitranjan Ranawat award functional outcome after total knee replacement varies with patient attributes. Clin Orthop Relat Res 2008;466(11):2597.
[19] Pelt CE, Anderson MB, Erickson JA, Gililland JM, Peters CL. Adding value to total joint arthroplasty care in an academic environment: the Utah experience. J Arthroplasty 2018;33(6):1636.
[20] 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.
[21] Nwachukwu BU, O'Donnell E, McLawhorn AS, Cross MB. Episode of care payments in total joint arthroplasty and cost minimization strategies. HSS J 2016;12(1):91.
[22] Jette DU, Stilphen M, Ranganathan VK, Passek SD, Frost FS, Jette AM. Validity of the AM-PAC "6-Clicks" inpatient daily activity and basic mobility short forms. Phys Ther 2014;94(3):379.
[23] Rose M, Bjorner JB, Becker J, Fries JF, Ware JE. Evaluation of a preliminary physical function item bank supported the expected advantages of the patient-reported outcomes measurement information system (PROMIS). J Clin Epidemiol 2008;61(1):17.
[24] Hung M, Clegg DO, Greene T, Saltzman CL. Evaluation of the PROMIS physical function item bank in orthopaedic patients. J Orthop Res 2011;29(6):947.
[25] Budtz-Jorgensen E, Keiding N, Grandjean P, Weihe P. Confounder selection in environmental epidemiology: assessment of health effects of prenatal mercury exposure. Ann Epidemiol 2007;17(1):27.
[26] Rabe-Hesketh S, Everitt B. A handbook of statistical analyses using Stata. New York: Chapman & Hall/CRC; 2003.
[27] Rabe-Hesketh S, Skrondal A, Pickles A. Reliable estimation of generalized linear mixed models using adaptive quadrature. Stata J 2002;2(1):1.
[28] Enders CK. Analyzing longitudinal data with missing values. Rehabil Psychol 2011;56(4):267.
[29] Mazza GL, Enders CK, Ruehlman LS. Addressing item-level missing data: a comparison of proration and full information maximum likelihood estimation. Multivariate Behav Res 2015;50(5):504.
[30] Manning WG, Basu A, Mullahy J. Generalized modeling approaches to risk adjustment of skewed outcomes data. J Health Econ 2005;24(3):465.
[31] Sibia US, Mandelblatt AE, Callanan MA, MacDonald JH, King PJ. Incidence, risk factors, and costs for hospital returns after total joint arthroplasties. J Arthroplasty 2017;32(2):381.
[32] Molloy IB, Martin BI, Moschetti WE, Jevsevar DS. Effects of the length of stay on the cost of total knee and total hip arthroplasty from 2002 to 2013. J Bone Joint Surg Am 2017;99(5):402.
[33] Ghomrawi HMK, Schackman BR, Mushlin AI. Appropriateness criteria and elective procedures — total joint arthroplasty. N Engl J Med 2012;367(26):2467.
[34] Katz JN, Winter AR, Hawker G. Measures of the appropriateness of elective orthopaedic joint and spine procedures. J Bone Joint Surg Am 2017;99(4):e15.
[35] Bae CJ, Griffith S, Fan Y, et al. The challenges of data quality evaluation in a joint data Warehouse. EGEMS (Wash DC) 2015;3(1):1125.