LEWIN GROUP®

CMS Comprehensive Care for Joint Replacement Model: Performance Year 4 Evaluation Report

Fourth Annual Report

HEALTH CARE AND HUMAN SERVICES POLICY, RESEARCH, AND ANALYTICS — WITH REAL-WORLD PERSPECTIVE.

Prepared for: **Centers for Medicare & Medicaid Services**

Submitted by: **The Lewin Group, Inc. with our partners: Abt Associates, GDIT, and Telligen**

September 2021

CMS Comprehensive Care for Joint Replacement (CJR) Model: Performance Year 4 Evaluation Report

Fourth Annual Report

Prepared for:

Jessica McNeely, Research and Rapid-Cycle Evaluation Group (RREG), Center for Medicare & Medicaid Innovation (CMMI), Centers for Medicare & Medicaid Services (CMS)

Submitted by:

The Lewin Group, Inc.

September 2021

Table of Contents

List of Appendices

Executive Summary

The Comprehensive Care for Joint Replacement (CJR) model tests whether episode-based payment and quality measurement for lower extremity joint replacements (LEJR) can lower payments and improve quality.[1] Implemented on April 1, 2016 by the Centers for Medicare & Medicaid Services' (CMS) Innovation Center, this mandatory model is an important component of CMS' strategy to use alternative payment models (APMs) to slow Medicare spending growth by rewarding value rather than volume.[2]

The fourth annual CJR model evaluation report presents findings from the first four performance years, which include episodes initiated on or after April 1, 2016 that ended by December 31, 2019. At the start of performance year (PY) 3, the number of mandatory Metropolitan Statistical Areas (MSAs) was scaled back from the 67 original randomly selected MSAs to the 34 MSAs with the highest average historical payments.[3] This report focuses on the 395 mandatory CJR hospitals in the 34 mandatory MSAs that were continuously required to participate through the entire model; it also provides information on 74 opt-in and 200 non-opt-in hospitals in the 33 voluntary MSAs where hospitals could elect to continue participating in the CJR model. The analyses presented in this report do not include the CJR hospitals located in mandatory MSAs that were designated as low-volume or rural, which had a choice to voluntarily continue in the CJR model after the second performance year. Of the 126 low-volume or rural hospitals in mandatory MSAs, 15 chose to continue participating in the CJR model and 111 did not.

In the first four performance years, mandatory CJR hospitals achieved a statistically significant reduction in average episode payments due to reductions in institutional post-acute care (PAC) use. After accounting for reconciliation payments made to mandatory CJR hospitals, the payment reductions made by mandatory CJR hospitals likely resulted in Medicare savings during the first four performance years, although we cannot conclude with statistical certainty that the CJR model resulted in savings because there is a wide range around the estimated savings. Quality of care, as measured by the unplanned readmission rate, emergency department (ED) use, and mortality, improved or was maintained under the CJR model. Hospitals' relationships with orthopedic

---

[1]   The term LEJR refers to all discharges under Medicare Severity-Diagnosis Related Groups 469: Major Joint Replacement or Reattachment of Lower Extremity with major complications and comorbidities and 470: Major Joint Replacement or Reattachment of Lower Extremity without major complications and comorbidities. Starting in 2018, Medicare began covering total knee arthroplasty (TKA) procedures performed in a hospital outpatient department (Current Procedural Terminology (CPT) code 27447). Outpatient TKAs will be episodes under the CJR model beginning in October 2021. Appendix A includes an acronym list and glossary for terms used through this report.

[2]   Press MJ, Rajkumar R, Conway PH. Medicare's new bundled payments: design, strategy, and evolution [published online December 17, 2015]. JAMA. doi:10.1001/jama.2015.18161.

[3]   MSAs are counties associated with a core urban area that has a population of at least 50,000. Non-MSA counties (no urban core area or urban core area of less than 50,000 population) and MSAs with a volume of LEJR cases below 400 were not eligible for selection. Hospitals are required to participate in the CJR model if they are acute care hospitals actively engaged in Medicare and paid under the Inpatient Prospective Payment System. Hospitals are excluded if they are currently participating in a Bundled Payments for Care Improvement model for LEJR or are cancer hospitals.

surgeons and PAC providers shed light on how hospitals achieved the goals of the CJR model. Surveyed orthopedic surgeons indicated that hospitals provided guidelines or directives, shared performance feedback or data, or engaged surgeons in a financial gainsharing arrangement. The amount of care redesign direction that PAC provider representatives reported receiving from hospitals varied by provider type. Skilled nursing facility (SNF) representatives we interviewed reported changing care for LEJR patients in response to hospitals' requests, while outpatient physical therapists we interviewed reported that hospitals did not request changes to care plans for LEJR patients.

A. Introduction

CJR participant hospitals are accountable for the cost and quality of health care services for LEJR episodes of care. LEJR surgeries are primarily hip replacements (total hip arthroplasty or THA) and knee replacements (total knee arthroplasty or TKA). An episode of care begins with the hospitalization for the LEJR surgery and extends through the 90 days after hospital discharge. The CJR model financially rewards participant hospitals for reducing episode payments and improving quality, which hospitals may achieve by coordinating care with the surgeons, PAC providers, clinicians, and other providers involved in the episode. Through an annual reconciliation process, participant hospitals may earn a supplemental payment from Medicare for achieving cost and quality metrics or face repayments to Medicare if they do not.

The CJR model originally required hospitals in 67 MSAs to participate. Because of the CJR model's mandatory and randomized design, a spectrum of hospitals with varying levels of infrastructure, care redesign experience, episode costs, utilization, and market positions participated, which allowed a broad test of the CJR model. In the third performance year, beginning January 2018, CMS scaled back the number of mandatory MSAs to the 34 MSAs with the highest average historical episode payments. Hospitals in these mandatory MSAs that were not designated as low-volume or rural (mandatory CJR hospitals) were required to continue their participation in the CJR model.[4] Hospitals in the 33 MSAs with lower average historical payments (voluntary MSAs) and all hospitals that were designated as low-volume or rural had a one-time opportunity to opt-in to the CJR model for PY3-5. This report focuses on the 395 mandatory CJR hospitals. It also includes information about the impact of the CJR model on payments, utilization, quality of care, and savings to Medicare for 74 opt-in hospitals and 200 non-opt-in hospitals in the voluntary MSAs. The analyses in this report do not include the CJR hospitals located in mandatory MSAs that were designated as low-volume or rural and therefore had a choice to voluntarily continue in the CJR model for PY3-5. Of the 126 low-volume or rural hospitals in mandatory MSAs, 15 chose to continue participating in the CJR model and 111 did not.

---

[4] Low volume was defined as less than 20 episodes over the three-year historical baseline period (episodes that begin in 2012-2014). Rural was defined based on the Inpatient Prospective Payment System (IPPS) §412.103 rural reclassification list (as of January 31, 2018) or location within a rural Census tract of a MSA as determined by the Office of Rural Health Policy (ORHP) of the Health Resources and Services Administration.

Additional detail about the CJR model is in the Third Annual Evaluation Report, available at https://innovation.cms.gov/data-and-reports/2020/cjr-thirdannrpt, and the Center for Medicare & Medicaid Innovation (CMMI) webpage, https://innovation.cms.gov/initiatives/cjr.

This evaluation draws from a range of data sources, including Medicare claims, Medicare patient assessments, patient surveys, an orthopedic surgeon survey, telephone interviews, and program information, as well as various research methods, to understand the impact of the CJR model. Our evaluation examines the extent to which participant hospitals decide if and how to respond to the model, and how hospitals' decisions reflect hospital resources and market conditions. The impact of the CJR model is influenced by those decisions, as well as the relationship between a hospital's average historical episode payment relative to its quality-adjusted target price, and the type and magnitude of care redesign needed to earn reconciliation payments or avoid repayments. The evaluation approach provides insights into the relative successes and challenges in reducing episode payments and improving quality, and provides evidence on how hospitals in a variety of circumstances responded.

B. Results

The CJR model resulted in decreases in average payments (standardized allowed amounts) for LEJR episodes at mandatory CJR hospitals during the first four performance years. Average payments for LEJR episodes decreased by $1,511 more than for the control group (Appendix C, Section II details how the control group was defined). This relative reduction in payments equates to a 5.2% decrease in average episode payments from the baseline ($p<0.01$).[5]

> *Baseline* refers to a three-year period (2012 through 2014) prior to the CJR model that serves as a benchmark against which performance under the CJR model is compared.
>
> *Intervention* refers to the period that the CJR model was in effect and studied through our analyses. Except where noted, the intervention period includes the first four performance years of the CJR model (April 2016 through December 2019).

The statistically significant decrease in payments was primarily due to reductions in institutional PAC use. Mandatory CJR hospitals discharged relatively fewer patients to an inpatient rehabilitation facility (IRF) and CJR patients discharged to a SNF spent relatively fewer days in that care setting. Additionally, more CJR patients were first discharged to a home health agency (HHA), generally a lower cost PAC setting than IRF or SNF, than patients from control group hospitals. These changes in utilization resulted in statistically significant decreases in IRF and SNF payments, which drove the decrease in average episode payments.

Medicare achieves savings under the CJR model when reductions in episode payments are greater than the net reconciliation payments made to hospitals. Across the first four performance years, Medicare likely realized savings from mandatory CJR hospitals, but we cannot conclude with statistical certainty that there was savings due to the wide range around the estimated savings. Estimated Medicare savings was $76 million (ranging from possible losses of $15.3

---

[5] This value represents the percent change from the CJR baseline that is due to the CJR model. It is calculated by dividing the difference-in-differences (DiD) estimate by the CJR baseline average.

million to savings $167.2 million) from mandatory CJR hospitals during the first four performance years. Medicare savings is presented as a range because it is based on our statistical analysis of the reduction in payments, which includes a range that is intended to capture uncertainty around our estimate.

Quality of care, as measured by the unplanned readmission rate, emergency department use, and mortality, improved or was maintained for mandatory CJR hospitals. For patients who were first discharged from a mandatory CJR hospital to an IRF, SNF, or HHA, improvement in functional status and pain during their PAC stay was generally similar for CJR and control group patients. By the end of the episode, on average, CJR and control survey respondents reported similar functional status gains and pain levels from before the episode. CJR and control survey respondents reported similar satisfaction with overall recovery and care management and similar care transition experiences, although CJR survey respondents required more help from caregivers when they returned home compared to control respondents. Orthopedic surgeons, other clinicians, and hospital staff we interviewed or consulted indicated that shifting recovery to the home setting could increase caregiver responsibilities, but were consistent in their view that home was the best place for most patients to recover, congruous with the research base.[6,7]

The survey results differed for patients with a hip fracture, who are typically more severe in terms of decline in physical function and length of recovery.[8,9] Among survey respondents with a hip fracture, CJR respondents reported worse functional status and required more caregiver help than control respondents. Specifically, CJR respondents reported less improvement in three of eight measures: rising from sitting, standing, and using the toilet. For each of these three measures, more than 60% of respondents with a hip fracture in both the CJR and control groups regained or exceeded their pre-hospital function. However, for each measure, the differences between the CJR and control groups equate to roughly three to six more CJR respondents out of every 100 CJR respondents declining in function from before their fracture until after the end of the episode. There were no statistically significant differences in the other five measures of functional status, although CJR respondents had point estimates indicating less improvement in four of those five remaining measures. The difference between the CJR and control groups in caregiver help equates to approximately 5 to 10 more CJR respondents out of 100 requiring more caregiver help. Despite differences in functional status and the need for caregiver help, there were no differences in

[6]   Keswani, A., BA, Tasi, M. C., BA, Fields, A., BA, Lovy, Andrew J., MD, MS, Moucha, C. S., MD, & Bozic, Kevin J., MD, MBA. (2016). Discharge destination after total joint arthroplasty: An analysis of postdischarge outcomes, placement risk factors, and recent trends. *Journal of Arthroplasty, The*, 31(6), 1155-1162. doi:10.1016/j.arth.2015.11.044

[7]   Slover, James, M.D., M.S. (2016). You want a successful bundle: What about post-discharge care? *Journal of Arthroplasty, The, 31*(5), 936-937. doi:10.1016/j.arth.2016.01.056

[8]   Le Manach Y, Collins G, Bhandari M, et al. Outcomes After Hip Fracture Surgery Compared With Elective Total Hip Replacement. *JAMA.* 2015;314(11):1159–1166. doi:10.1001/jama.2015.10842

[9]   Wu VJ, Ross BJ, Sanchez FL, Billings CR, Sherman WF. Complications Following Total Hip Arthroplasty: A Nationwide Database Study Comparing Elective vs Hip Fracture Cases. J Arthroplasty. 2020 Aug;35(8):2144-2148.e3. doi: 10.1016/j.arth.2020.03.006. Epub 2020 Mar 10. PMID: 32229152.

patients' reported satisfaction with overall recovery. These are important findings that warrant additional investigation.

The evaluation also examined whether the CJR model resulted in any unintended consequences. For example, CJR participant hospitals could increase the volume of LEJRs to maximize reconciliation payments, delay services until after the end of the episode to keep episode payments low, or favor less complex patients who may be less costly to treat. The CJR model was designed to protect against these responses by including all hospitals in the MSA, using a long episode definition, and other means, but the CJR model may still result in unintended consequences. We found no evidence that the CJR model was associated with an increase in the total market volume of LEJRs. It is unlikely that mandatory CJR hospitals shifted services to after the end of the episode. Mandatory CJR hospitals did experience a decrease in patient complexity relative to control group hospitals for elective LEJRs without major complications or comorbidities, which contributed to the reduction in average episode payments.

The change from mandatory to voluntary status for about half of the MSAs in the third performance year provided an opportunity to evaluate previously mandatory participants that opt-in to a voluntary APM. In the third annual evaluation report, we found that hospitals that opt-in likely did so because they anticipated receiving reconciliation payments. In this report, we find opt-in hospitals received per episode reconciliation payments that were much higher than their payment reductions, resulting in statistically significant losses for Medicare. The voluntary component of the CJR model also provided an opportunity to evaluate previous participants that do not opt-in to a voluntary APM. The episode payment reductions for non-opt-in hospitals when they were mandatory participants in the CJR model did not persist after they stopped participating in the CJR model, which raises doubts about lasting effects of a mandatory APM after participation ends. It is important to note, however, that this result may not be generalizable to other APMs because non-opt-in hospitals were located in MSAs with lower average historical episode payments and chose not to continue their participation in the CJR model.

The relationships between CJR participant hospitals and orthopedic surgeons and PAC providers are important to consider in the context of the evaluation; while the CJR model holds hospitals accountable for LEJR episodes, orthopedic surgeons and PAC providers influence episode costs through the decisions they make and services they provide. Surgeon survey respondents indicated that CJR participant hospitals tried to influence their decisions under the model through a variety of means including by providing guidelines or directives to consider when determining whether to perform an LEJR, performance feedback or data, and financial gainsharing arrangements. PAC provider representatives we interviewed provided varying indications that hospitals aimed to influence care redesign after discharge from the hospital. Outpatient physical therapists (PTs) generally did not indicate that hospitals requested changes to LEJR patient care plans, while SNF administrators indicated that hospitals requested changes, such as increased information exchange, reduced length of stay, or adjusted physical therapy schedules for LEJR

patients. Both PTs and SNF administrators indicated that their operations or care provision were influenced by factors including but not limited to the CJR model.

These key findings build on our evaluation framework, which focuses on the hospital where the LEJR episode begins and recognizes the influence of financial, hospital, and market factors on hospitals' responses to the CJR model. Accordingly, the research questions considered under this evaluation were organized into three categories: (1) impact of the CJR model, (2) financial risk or opportunity, and (3) CJR participant hospital responses. Additional details about key findings are summarized under the main research questions addressed in this report.

1. Impact of the model

a. What was the impact of the CJR model on average episode payments?

- **For mandatory CJR hospitals, there were statistically significant reductions in average episode payments during the first four performance years of the CJR model.** [10] Average payments for LEJR episodes decreased by $1,511 more for mandatory CJR hospitals than for the control group. This relative reduction in payments equates to a 5.2% decrease in average episode payments from the baseline (p<0.01).[11]

b. How much did the Medicare program save or lose due to the CJR model after accounting for reconciliation payments?

- **Across the first four performance years, Medicare likely realized savings from mandatory CJR hospitals, although we cannot conclude this with certainty due to the wide range around our savings estimate.** Estimated savings were $76.0 million, ranging from losses of $15.3 million to savings of $167.2 million. The savings estimate is based on an estimated reduction in total non-standardized paid amounts of $202.0 million less net reconciliation payments of $126.1 million.[12] These estimates are equivalent to savings of $494 per episode (ranging from losses of $99 per episode to savings of $1,087 per episode), and equate to savings of 1.82% from the baseline for mandatory CJR hospitals. Medicare savings estimates were lower in PY3 and PY4 because mandatory CJR hospitals shifted fewer TKAs to the lower payment outpatient setting.

- **Considering opt-in hospitals for the first four performance years, and non-opt-in hospitals for the first two performance years, we cannot conclude that Medicare realized savings across the entire CJR model.** Estimated losses from opt-in hospitals were $44.5 million, ranging from losses of $62.9 million to losses of $26.1 million. Estimated losses from non-opt-in hospitals were $10.0 million, ranging from losses of $20.2 million to savings of $0.2 million. Across all hospitals, estimated Medicare

---

[10] Episode payments are defined as Medicare standardized allowed amounts. Standardization removes the effect of wage and other Medicare payment adjustments. Allowed amounts include beneficiary cost sharing.

[11] This value represents the percent change from the CJR baseline that is due to the CJR model. It is calculated by dividing the DiD estimate by the CJR baseline average.

[12] The $126.1 million is based on $179.1 million in reconciliation payments made to mandatory CJR hospitals less $53 million in repayments received from mandatory CJR hospitals.

savings were $21.4 million, ranging from losses of $75.0 million to savings of $117.9 million. Due to this wide range, which includes the possibility of substantial losses or savings, we cannot conclude that the CJR model resulted in savings to Medicare across all hospitals that ever participated.

c.   What was the impact of the CJR model on service-level payments and service use?

- **The relative decrease in average episode payments was driven by reductions in the use of institutional PAC.** Mandatory CJR hospitals discharged relatively fewer patients to an IRF (a 28.1% decrease from the CJR baseline proportion, $p<0.01$), which resulted in a relative decrease in IRF payments of $593 per episode ($p<0.01$). Patients who received LEJRs at mandatory CJR hospitals and had a SNF stay spent an average of 2.6 fewer days in a SNF relative to control patients ($p<0.01$), which contributed to a relative decrease in SNF payments of $843 per episode ($p<0.01$). Additionally, more CJR patients were first discharged to an HHA (a 20.5% increase from the CJR baseline proportion, $p<0.05$).

d.   What was the impact of the CJR model on quality of care?

- **Quality of care measures improved or were maintained under the CJR model.** During the first four performance years, the unplanned readmission rate decreased more for CJR episodes than for control group episodes, representing a 3.5% decrease from the CJR baseline ($p<0.10$). For elective LEJR episodes, there was a relative reduction in the complication rate, representing a 7.9% decrease from the CJR baseline ($p<0.05$). There were no statistically significant changes in emergency department use or mortality.

e.   What was the impact of the CJR model on functional status, pain, and care experiences?

- **In mandatory MSAs, CJR and control patients who first received IRF, SNF, or HHA care after their LEJRs showed similar improvement during the PAC stay.** These beneficiaries had similar short-term changes in functional status, measured as improvement in functional status and pain from the first to the last patient assessment completed during the PAC stay.

- **For the subset of survey respondents with a hip fracture, CJR respondents reported less improvement in functional status from before their LEJR to the end of their episode than control respondents.** When considering all survey respondents (patients with elective LEJR and LEJR due to hip fracture), CJR and control survey respondents had similar self-reported changes in functional status and pain from before their LEJR to the end of the episode.

- **CJR and control survey respondents generally reported similar satisfaction with overall recovery and care management and similar experience with care transitions.**

- **CJR survey respondents required more help from caregivers when they returned home compared to control respondents.** The greater need for caregiver help was more

pronounced for the subset of CJR survey respondents with a hip fracture relative to control group respondents with hip fracture.

f.   Did the CJR model result in any unintended consequences?

What was the impact of the CJR model on total market volume of elective LEJR discharges?

- **The CJR model had no statistically significant impact on the total volume of elective LEJR discharges (inpatient plus outpatient).** The estimated impact of the CJR model on market-level LEJR discharge rates was a decrease of 0.08 discharges per 1,000 fee-for-service (FFS) beneficiaries, an estimate that is small and not statistically significant (p=0.21).

Were there any indications that the CJR patient population was different in the intervention period than in the baseline period?

- **Mandatory CJR hospitals had a decrease in patient complexity relative to control group hospitals for elective LEJRs without major complications or comorbidities.** Less complex patients may require fewer resources and, therefore, have less costly episodes. Thus, a relative reduction in patient complexity could make it easier for CJR hospitals to receive reconciliation payments without improving the efficiency of how they provide care during the episode or improving the quality of care. During the first four performance years, mandatory CJR hospitals reduced average episode payments by $170 per episode relative to control hospitals because their average patient complexity decreased relative to the average complexity of control patients. Our estimated impacts of the CJR model on average episode payments discussed earlier do not include this change in payments resulting from relative changes in patient mix because they are intended to measure the impact of the CJR model on episode payments for similar patients. Therefore, the $170 relative reduction in payments resulting from relative reductions in the complexity of CJR patients is in addition to the overall impact estimate. This finding suggests that some of the reconciliation payments made to CJR hospitals may be due to their decrease in patient complexity, which was not accounted for in the reconciliation payment calculation.

- **For the more complex episode groups – elective LEJR with major complications or comorbidities and LEJRs due to fracture – the CJR patient population was not relatively healthier than control patients.**

What was the impact of the CJR model on payments in the 30 days following the episode?

- **Payments for services provided during the 30 days following the episode were likely not affected by the CJR model.** Our estimate indicates that during the first four performance years the CJR model had no statistically significant impact on payments for services provided during the 30 days following the episode, which we monitor for any indication that services were postponed until after the end of the episode. We cannot be certain, however, that there was no impact of the CJR model on this outcome because it failed parallel trends tests. In other words, CJR and control group hospitals may have had

different patterns of post-episode payments in the baseline, which makes it difficult to isolate the impact of the CJR model on this outcome.

g.   What is the impact of the CJR model on health equity?

- **For mandatory CJR hospitals, there is limited evidence of different impacts of the CJR model on patient populations with historically poorer access to care and health outcomes.**[13] We studied changes in payments, quality, functional status, satisfaction, and caregiver help for three subpopulations of patients: those who are Black or African American, those eligible for both Medicare and Medicaid (dually eligible), and those who are both Black or African American and dually eligible. The CJR model resulted in larger payment reductions for patients who were Black or African American than for white patients. We found limited evidence of different impacts of the CJR model on quality of care (as measured by emergency department use, readmissions, and mortality), functional status, satisfaction, and caregiver help. For the large majority of these outcomes, estimated differential impacts were statistically insignificant, indicating no systematic differences in these measures of quality of care. For a few outcomes that did indicate a statistically significant differential impact, the CJR model improved the quality of care for the subpopulation of study. It is important to note, however, that these results do not account for potential changes in LEJR volume or the complexity of patients receiving LEJR within a subpopulation. Our patient mix analyses have consistently indicated that there was a relative reduction in the proportion of dually eligible patients at mandatory CJR hospitals and in the past we also observed a reduction in the proportion of CJR patients who were Black or African American. It is possible that any relative improvements for a given subpopulation could be a reflection of changes in the patient population, rather than actual quality improvements due to the CJR model. Thus, the current results need to be interpreted with caution and changes in LEJR volume and patient mix in subpopulations warrants further examination.

h.   What was the impact of the CJR model on hospitals in voluntary MSAs?

- **For opt-in CJR hospitals, there was a statistically significant reduction in average episode payments during the first four performance years of the CJR model.** Average payments decreased by $752 more than for matched control hospitals, which equates to a 3.2% decrease from the CJR baseline. The reduction in average episode payments was mainly driven by a relative reduction in SNF payments.

- **For opt-in CJR hospitals, changes in quality of care were mixed.** During the first four performance years, there was a relative reduction in ED use of approximately 5.9% and no change in the unplanned readmission rate or mortality. Patients discharged from opt-in CJR hospitals to an IRF, SNF, or HHA had less improvement in functional status while in

---

13   U.S. Department of Health and Human Services Office of the Assistant Secretary for Planning and Evaluation. (2016). *Report to Congress: Social Risk Factors and Performance Under Medicare's Value-Based Purchasing Programs.* https://aspe.hhs.gov/system/files/pdf/253971/ASPESESRTCfull.pdf

the PAC setting than patients at matched control group hospitals, which could signal potential quality concerns.

- **For non-opt-in CJR hospitals, there was a statistically significant reduction in average episode payments during the first two performance years, but this impact did not persist after non-opt-in CJR hospitals stopped participating in the CJR model.** When non-opt-in hospitals participated in the CJR model (PY1-2), average episode payments decreased by $440 relative to the matched control group hospitals (1.7% decrease from the baseline). After they stopped participating (PY3-4), there was no longer a statistically significant impact on episode payments. Consistent with the changes in payments, there are indications that, after they stopped participating in the CJR model, non-opt-in hospitals reduced the proportion of patients discharged to an HHA, a lower payment setting, and increased the proportion of patients discharged to a SNF, a higher payment setting. Conversely, reductions in the length of SNF stays persisted after the hospitals stopped participating in the CJR model, which could be an indication of a more lasting impact.

- **Patients discharged from non-opt-in CJR hospitals to a SNF or HHA during the first two performance years exhibited less improvement in functional status during their PAC stay relative to patients at matched control group hospitals.**

2. *Financial risk or opportunity*

a. *What factors were associated with receiving reconciliation payments under the CJR model?*

- **Half of mandatory CJR hospitals received reconciliation payments in PY3 and PY4, when the target price was based largely or fully on the regional average.** In contrast, 58% and 69% of mandatory hospitals received reconciliation payments in PY1 and PY2, respectively, when the target price was more heavily weighted to hospital-specific historical payments. As the target price shifted from being largely based on hospital-specific historical payments to being based on regional average payments, the target price decreased for most hospitals. The target price decreased from PY2 to PY3 for 91% of hospitals. For those hospitals, the average decrease was 7%. From PY3 to PY4, the target price decreased for 75% of mandatory hospitals, and for those hospitals, the average decrease was 4%. As a result, hospitals needed to reduce their episode payments to a greater degree in later years of the model to come below the target price and receive reconciliation payments.

- **The majority of hospitals received reconciliation payments under this mandatory model, but hospitals with certain characteristics received higher reconciliation payments.** In the first four performance years, approximately one-quarter of mandatory CJR hospitals had average historical payments below their target price at the start of the performance year and these hospitals received higher reconciliation payments. These hospitals may not have needed to change their practice patterns to receive reconciliation payments under the CJR model. Additionally, the average reconciliation payment per

episode was higher for hospitals that served less complex patient populations, had higher composite quality scores, or had higher LEJR volume.

3. *CJR Participant Hospital Responses*

a. *How did the CJR model influence relationships between CJR participant hospitals and orthopedic surgeons?*

- **Orthopedic surgeon survey respondents indicated that CJR hospitals influenced clinical practice through, for example, providing guidelines or directives, sharing performance feedback, or implementing financial gainsharing agreements.** The relationships between hospitals and surgeons are important to consider in the context of the CJR model; while hospitals are accountable for LEJR episodes, orthopedic surgeons influence episode costs through their clinical decisions. Orthopedic surgeons responding to a survey indicated that hospitals provided guidelines or directives under the CJR model that surgeons should consider patient risk factors, particularly modifiable factors such as uncontrolled diabetes or obesity, when determining whether to perform an LEJR. Surgeon survey respondents reported that they often received performance feedback or data from hospitals, and this information influenced most of them to modify their care practices.

b. *How did the CJR model influence relationships between CJR participant hospitals and post-acute care providers?*

- **Post-acute care provider representatives reported varying care redesign direction from hospitals.** Consistent with the reduction in SNF payments under the CJR model, hospital respondents previously indicated that reducing SNF care was among the key objectives of their redesign activities and many indicated they leveraged relationships with post-acute care providers to influence care after patients were discharged from the hospital. SNF interviewees said they changed LEJR patient care in response to hospitals' requests. Most often, SNF administrators indicated that hospitals wanted them to provide information about the patient while at the SNF, reduce SNF length of stay, and adjust the frequency and timing of physical therapy. In contrast, most outpatient PTs we interviewed indicated that hospitals did not request changes to the care plans for their LEJR patients. Both SNF administrators and PTs indicated that their facility's operations and the care they provided to LEJR patients were influenced by a variety of factors, including the CJR model.

C. Discussion

This fourth annual evaluation report demonstrates that the CJR model, which holds hospitals accountable for payments and quality for an episode of care that begins with LEJR surgery, remains a promising approach for reducing episode payments. Through the fourth year of the model, participating hospitals continued to respond to its financial incentives by shifting patients to less intensive PAC settings, resulting in a relative reduction in episode payments. Overall quality of care was maintained or improved for mandatory hospitals, although CJR patients with a hip fracture reported less improvement in functional status. Additionally, CJR patients, particularly

those with a hip fracture, required more caregiver help. For patient populations with historically poorer access to care and health outcomes, there was limited evidence of different impacts of the CJR model on quality of care, functional status, satisfaction, and caregiver help. We will continue to monitor the impact of the CJR model on patients with fractures and subpopulations with historically poorer access to care and health outcomes. Nevertheless, for the majority of patients, the CJR model reduced episode payments without compromising quality of care.

The reductions in episode payments likely resulted in Medicare savings for mandatory hospitals. Medicare savings was reduced in the third and fourth performance years, however, because mandatory hospitals shifted fewer TKAs to the hospital outpatient setting than the control group. Furthermore, there are indications that some hospitals may have received reconciliation payments at least in part because they treated a healthier mix of patients. As in prior years, there continues to be evidence that the simple risk stratification methodology based on MS-DRG and fracture status used to set quality-adjusted target prices did not adequately account for variations in patient complexity that affected episode payments, which likely dampened Medicare savings. Beginning in PY6, the quality-adjusted target prices will include more comprehensive risk adjustment that may, in turn, affect Medicare savings under the CJR model. While Medicare likely realized savings from mandatory hospitals, Medicare did not realize savings for all hospitals that ever participated in the CJR model. Hospitals in the 33 voluntary MSAs that chose to continue participating in the model contributed to Medicare losses, and hospitals that did not continue participating likely also contributed to Medicare losses. Notably, hospitals that stopped participating in the CJR model after the second performance year did not continue to reduce average episode payments. This novel finding raises questions about the permanence of the effects of mandatory APMs.

In future reports, we will deepen our understanding of the impact of the CJR model by refining our estimates of Medicare program savings while adjusting for other policies that affect service use and payments, further investigating the relationship between target prices and savings, and incorporating subgroup analyses. We will continue to monitor for unintended consequences and further explore the variation in patient complexity and its impact on reconciliation payments and Medicare savings. As the payment landscape continues to evolve, and as hospitals respond to the COVID-19 pandemic, we will continue to evaluate how broader Medicare policy changes and the evolving health care delivery landscape affect the impact of the CJR model.

I. Introduction

The CJR model tests whether episode-based payments and quality measurement for LEJR can lower payments and improve quality.[14] Implemented on April 1, 2016 by the CMS Innovation Center, this mandatory model is an important component of CMS' strategy to use APMs to slow Medicare spending growth by rewarding value rather than volume.[15]

The fourth annual CJR model evaluation report presents findings from the first four performance years, which include episodes initiated on or after April 1, 2016 that ended by December 31, 2019. During the first two years, the model was mandatory for nearly all acute care hospitals in 67 geographic areas, defined by MSAs.[16] At the start of PY3, the number of mandatory MSAs was scaled back to the 34 MSAs with the highest historical payments; hospitals in the other 33 MSAs were given the opportunity to continue to participate voluntarily. This report primarily focuses on the hospitals in the 34 MSAs that were required to participate in the CJR model throughout the intervention period (mandatory CJR hospitals).

A. The CJR Model

CJR participant hospitals are financially accountable for the cost and quality of health care services for LEJR episodes of care. The CJR model rewards participant hospitals for reducing episode payments and improving quality, which provides hospitals with incentives to coordinate care with the physicians, PAC providers, and other providers and clinicians involved in the episode.[17] Through an annual reconciliation process, participant hospitals may earn reconciliation payments if they achieve cost and quality metrics or face repayments to Medicare if they do not. The CJR model originally required hospitals in 67 randomly selected MSAs to participate. Because of this mandatory, randomized design, a spectrum of hospitals with varying levels of infrastructure, care redesign experience, episode costs and utilization, and market positions participated, which allowed a broad test of the CJR model. In the third performance year, CMS changed the CJR

---

[14] The term LEJR refers to all discharges under Medicare Severity-Diagnosis Related Groups 469: Major Joint Replacement or Reattachment of Lower Extremity with major complications and comorbidities and 470: Major Joint Replacement or Reattachment of Lower Extremity without major complications and comorbidities. Appendix A includes an acronym list and glossary for terms used through this report.

[15] Press MJ, Rajkumar R, Conway PH. Medicare's new bundled payments: design, strategy, and evolution [published online December 17, 2015]. JAMA. doi:10.1001/jama.2015.18161.

[16] MSAs are counties associated with a core urban area that has a population of at least 50,000. Non-MSA counties (no urban core area or urban core area of less than 50,000 population) and MSAs with a volume of LEJR cases below 400 were not eligible for selection. Hospitals are required to participate in the CJR model if they are acute care hospitals actively engaged in Medicare and paid under the Inpatient Prospective Payment System. Hospitals are excluded if they are currently participating in a Bundled Payments for Care Improvement LEJR model or are cancer hospitals.

[17] The CJR model waives certain Medicare payment rules and fraud and abuse laws so participant hospitals have more flexibility to collaborate with clinicians and PAC providers. Appendix B includes more information about the CJR model waivers.

model so that only hospitals located in the 34 MSAs with the highest historical episode payments were required to continue their participation for the final three years of the model.

Episode definition. Under the CJR model, an LEJR episode of care begins with the hospitalization for the surgery and extends through the 90 days after hospital discharge. All Medicare-covered items and services provided during this period, with some exclusions, are in the episode bundle.[18] All providers and suppliers involved in the episode continue to be paid under Medicare's FFS payment systems.

Annual reconciliation. After the end of each model performance year, CMS reconciles each participant hospital's LEJR episode payments against the hospital's quality-adjusted target price. The quality-adjusted target price is based on a discounted blend of the hospital's average historical episode payments and the region's average historical episode payments. During the first two performance years, two-thirds of the quality-adjusted target price was the hospital's average and one-third was the regional average. In PY3, two-thirds of the quality-adjusted target price was the regional average and one-third was the hospital's average historical episode payment. Beginning in PY4, all quality-adjusted target prices were based on the regional average. The quality-adjusted target price is based on a rolling three-year historical period, and in PY3 and PY4 the historical period included the first year of the CJR model.

The discount to the quality-adjusted target price is intended to be Medicare's portion of the decrease in spending under the model. At reconciliation, the discount is adjusted based on the participant hospital's composite quality score. A lower discount is applied to the target price for participant hospitals with a higher quality score, thus rewarding higher quality through a higher quality-adjusted target price.

Hospitals with LEJR episode payments *below* their quality-adjusted target price and an "acceptable" or higher composite quality score receive a reconciliation payment. The reconciliation payment equals the difference between the quality-adjusted target price and actual episode payments, up to a stop-gain limit. Starting in PY2, hospitals with episode payments *above* their quality-adjusted target price repay Medicare the difference, subject to a stop-loss limit. In PY1, this repayment responsibility was forgiven to allow hospitals time to gain experience under the CJR model before implementation of two-sided risk. Both opportunity and risk increase over time as stop-gain and stop-loss limits increase. The stop-gain limit increased from 5% in PY1 to 20% in PY4 and the stop-loss limit increased from 0% to 20% over the same period.

---

[18]   Excluded items, services, and payments include: hemophilia clotting factors; new technology add-on payments; transitional pass-through payments for medical devices; items and services unrelated to the anchor hospitalization as specified by CMS on the CJR website, including (i) inpatient hospital admissions for oncology, trauma medical, chronic disease surgical, and acute disease surgical diagnoses, (ii) Medicare Part B services for acute disease and certain chronic disease diagnoses, (iii) certain per beneficiary per month payments; certain incentive programs and add-on payments under existing Medicare payment systems; and payments for otherwise included items and services in excess of two standard deviations above the mean regional episode payment.

Mandatory, randomized design of the original CJR model. The original mandatory, randomized design of the CJR model resulted in a diverse group of CJR participant hospitals, including hospitals that might not voluntarily participate in an episode-based payment model. For the first two performance years, all acute care hospitals paid under the Medicare Inpatient Prospective Payment System (IPPS), with few exceptions, in 67 randomly selected MSAs were required to participate. The original 67 mandatory MSAs were identified from 171 MSAs that were eligible for participation when the model design was finalized. MSAs were selected for participation using eight sampling strata based on a median split of MSA population size and quartiles of average MSA historical episode payments.[19] An MSA's probability of selection increased with the payment quartiles to oversample high-payment MSAs for participation in the CJR model. This was because of the belief that there is greater need and more opportunities for payment reductions in higher payment areas. Eligible MSAs that were not selected are a natural control group for evaluating the impact of the CJR model.

Changes to the CJR model in 2018. Effective January 2018, CMS scaled back the CJR model to the 34 MSAs with the highest historical episode payments (mandatory MSAs). Hospitals in these mandatory MSAs that were not designated as low-volume or rural were required to continue their participation in the CJR model. In January 2018 the number of hospitals ever-required to participate in the CJR model was reduced from 831 hospitals to 395 hospitals. CMS allowed the remaining hospitals in the 33 lower payment MSAs (voluntary MSAs) and all hospitals that were designated as low-volume or rural a one-time opportunity to opt-in to the CJR model for PY3-5. Of the 310 hospitals in the 33 voluntary MSAs, 75 opted to continue their participation in the model (opt-in hospitals) and 235 stopped participating in the model (non-opt-in hospitals) (Exhibit 1).

---

[19]   Originally, 196 MSAs were identified as eligible for participation in the CJR model and the mandatory MSAs were randomly selected from this pool. CMS later identified 25 MSAs that were ineligible for selection after accounting for Bundled Payments for Care Improvement (BPCI) physician group practice participation.

Exhibit 1: **This 4ᵗʰ annual report focuses on the 395 mandatory CJR hospitals and includes analyses of the two subgroups of hospitals located in the 33 voluntary MSAs**
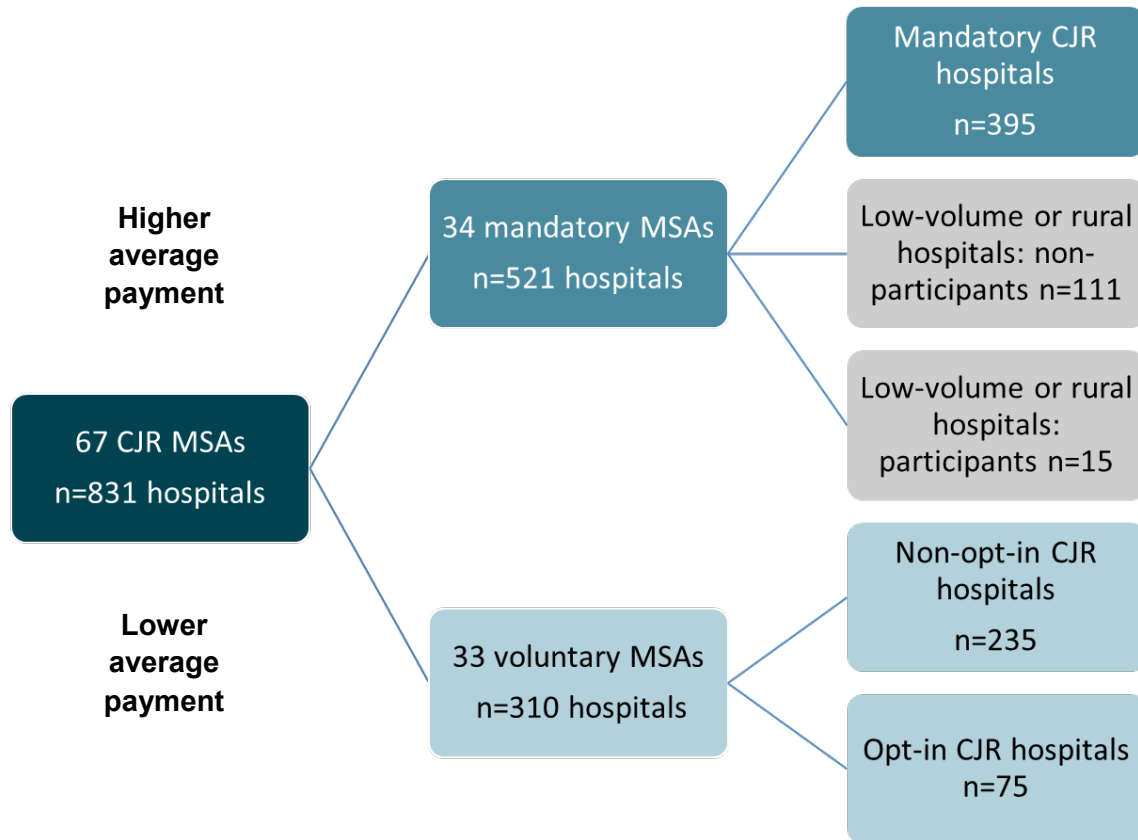
Notes: The numbers of hospitals in this exhibit include those that ever-participated in the CJR model.
          Gray boxes represent hospital groups not included in this annual report.
          MSA = metropolitan statistical area, PY = performance year.

The randomized design of the CJR model was not completely preserved when the model was scaled back in PY3 because the 34 mandatory MSAs were identified using a median split based on historical payments rather than the original sampling strata that categorized MSAs as high payment. Most, but not all, of the MSAs in the top two payment quartiles were categorized in the mandatory MSA group. The characteristics of mandatory CJR hospitals reflect their location in higher payment markets; however, the mandatory CJR hospitals remain a diverse group with variation in terms of LEJR volume, patient complexity, and institutional characteristics.[20] Hospitals in high-payment MSAs that were eligible but not selected for the CJR model serve as the control group for mandatory CJR hospitals.

Other policy changes in 2018. CMS implemented broad changes to Medicare coverage that affect the evaluation of the CJR model. The CJR model implemented episode-based payments for inpatient LEJR, which are primarily hip replacements (THA) and knee replacements (TKA). Effective January 2018, CMS removed TKA from Medicare's inpatient only list to allow Medicare

---

[20]   Centers for Medicare & Medicaid Services. Comprehensive care for joint replacement model - third annual report. https://innovation.cms.gov/data-and-reports/2020/cjr-thirdannrpt. 2020:21-31

coverage of TKAs performed in the hospital outpatient department. When TKA is performed on an inpatient, the hospital's payment is made under the IPPS and the surgery triggers an episode if the hospital is participating in the CJR model. When TKA is performed on an outpatient, the hospital's payment is made under the outpatient payment system and the surgery does not trigger an episode. Although this policy change was independent of the CJR model, CJR participant hospitals responded to it differently than non-CJR hospitals. CJR participant hospitals shifted a smaller proportion of TKA surgeries to the outpatient setting than control group hospitals. In the third and fourth annual reports we refined our methodology to account for this differential response. Section II.A.1 describes these changes in more detail.

Additional detail about the CJR model is available in the Third Annual Evaluation Report, available at https://innovation.cms.gov/data-and-reports/2020/cjr-thirdannrpt and the CMMI webpage, https://innovation.cms.gov/initiatives/cjr.

B.   Evaluation Conceptual Framework

The conceptual framework for the evaluation of the CJR model (Exhibit 2) reflects the fundamental features of the model and is informed by health services research literature, including evaluations of other episode-based payment approaches.[21] The evaluation framework focuses on the hospital where the LEJR episode begins because the hospital has the incentives to control payments and improve quality across the entire episode. The hospital's resources and market conditions will affect its decisions about whether and how to respond to the model.

This evaluation draws from a range of data sources, including claims, patient assessments, a patient survey, site visits, interviews, and program information, and relies on various research methods to understand the impact of the CJR model. Together, these provide insights into the relative successes and challenges in reducing episode payments and improving quality, and provides evidence on how hospitals in a variety of circumstances achieved these changes.

---

[21]   Maniya, O. Z., Mather, R. C., Attarian, D. E., Mistry, B., Chopra, A., Strickland, M., & Schulman, K. A. (2017). Modeling the Potential Economic Impact of the Medicare Comprehensive Care for Joint Replacement Episode-Based Payment Model. *The Journal of Arthroplasty*.

Exhibit 2:   Key research questions and domains are based on the evaluation conceptual framework

What was the impact of the CJR model?
Evaluated through analyses of Medicare claims and assessment data and a patient survey

- Total episode payments
- Savings to Medicare
- Service-level payments and service use
- Quality of care
- Functional status, pain, and care experiences

- Total market volume of elective LEJR discharges
- Patient population
- Payments in the 30 days following the episode
- Variations in impact

How did CJR participant hospitals choose to respond to the model?
Evaluated through analyses of interviews, site visits, and a surgeon survey

- Choice of response
- Care redesign

Financial risk or opportunity
Evaluated through analyses of interviews, site visits, and reconciliation data

- Reconciliation payments
- Perceptions of financial pressure

Resources and market conditions
Evaluated through analyses of interviews and site visits

- Relationships and resources

Note: LEJR = lower extremity joint replacement.

Impact of the Model. The CJR model is designed to affect episode payments, utilization, and quality outcomes. We use Medicare claims data to determine the impact of the model on Medicare payments (and associated utilization patterns) for LEJR episodes by examining the change in these outcomes relative to the change in the control group. Analyses of Medicare claims demonstrate the magnitude of payment changes due to the CJR model and the source of payment changes by type of service. Relative differences in utilization patterns between the treatment and control group provide further insights into how participant hospitals responded to the model.

Medicare claims reveal impacts of the CJR model on quality outcomes. Self-reported measures from a patient survey and patient assessment data collected from PAC settings provide information on functional status and pain. The cross-sectional patient survey analysis compares patients in CJR episodes with patients in control episodes, providing insights into the relationship between CJR participation and patient experience.

The claims analysis reveals whether the CJR model resulted in participant hospitals reducing episode payments as intended. However, additional analysis is needed to determine if participant hospitals responded to the model by increasing the volume of episodes. Even if there was a reduction in per episode payments, an increase in volume could increase total Medicare spending. We examine whether the change in volume of elective LEJR discharges differs between mandatory CJR MSAs and control MSAs.

Whether the model ultimately results in savings to the Medicare program also depends on Medicare reconciliation payments and repayments under the model. The impact of the CJR model on episode payments and volume of episodes, combined with reconciliation data, are used to estimate Medicare program savings for mandatory CJR hospitals. (Section II.A.1 examines the impact on average episode payments, Section II.A.2 examines savings to the Medicare program.)

Choice of response. Hospital leaders must consider multiple organizational factors, in addition to the potential for financial risk or opportunity, and internal and external resources, in making the business case for whether and how to respond to the CJR model. Orthopedic surgery is one of multiple service lines that compete for staff and other resources. The CJR model is one initiative that may or may not align with initiatives from other payers, state-specific policies, local labor markets, and other factors. (Section II.C.2 discusses hospital relationships with PAC providers, Section II.C.1 discusses hospital relationships with orthopedic surgeons).

Financial risk or opportunity. The distance between the quality-adjusted target price and episode payments varies for each hospital due to its historical average payments and the regional average. Hospitals with lower historical payments that are located in higher payment areas will likely be under the least financial pressure under the model and have the greatest opportunities to earn reconciliation payments. Hospitals in the opposite position, with higher historical payments that are located in lower payment areas will be under the most pressure to implement changes to avoid repaying CMS under the CJR model. The specific situation of each hospital will affect its ability to earn reconciliation payments and its responses to the model. Because of the changes in the calculation of the quality-adjusted target price, the amount of financial pressure on hospitals will shift over time. (Section II.B.1 explores the market, hospital, and patient characteristics associated with the amount of reconciliation payments.)

Resources and market conditions. A hospital's internal resources and market conditions will provide opportunities or constraints on its responses to the model. Hospitals with more capital and operational resources, such as dedicated care coordination staff or robust health information technology infrastructure, may be better situated to redesign care for LEJR episodes. Other hospital resources – such as leadership support, experience with episode-based payment or similar payment models, ownership of PAC providers, or employment of surgeons – may also affect their choices as well as their success in reducing payments below their quality-adjusted target price. Market conditions, such as the supply and characteristics of other providers involved in the episode, will affect how and whether hospitals garner support for delivering care more efficiently during the episode.

II. Results

A. Impact of the Model

1. *What was the impact of the CJR model on average episode payments?*

---

a. *Key Findings*

- For mandatory CJR hospitals, average payments (standardized allowed amounts) for LEJR episodes decreased by $1,511 more than for LEJR episodes at control group hospitals during the first four performance years. This equates to a 5.2% decrease from the baseline.

- The reduction in average payments for PY3 and PY4 was smaller than the reduction in PY2 because CJR hospitals performed fewer TKAs in the outpatient setting, which has lower Medicare payments, than control hospitals.

---

b. Methods

The analysis used a difference-in-differences (DiD) method to estimate the differential change in average Medicare standardized allowed amounts (payments) between the baseline (April 2012 through March 2015) and intervention period (April 2016 through December 2019, or the first four performance years of the CJR model) for all LEJR episodes initiated at mandatory CJR hospitals[22] relative to those initiated at control group hospitals. We used standardized payments to ensure that observed payment differences reflect

Acronyms	
BPCI	Bundled Payments for Care Improvement
DiD	difference-in-differences
LEJR	lower extremity joint replacement
MSA	metropolitan statistical area
PGP	physician group practice
pp	percentage point
PY	performance year
SNF	skilled nursing facility
THA	total hip arthroplasty
TKA	total knee arthroplasty

actual differences in billed services rather than Medicare payment policies. We used allowed amounts to eliminate variation in payments due to whether beneficiaries have met their deductible when they had the LEJR surgery. We used the DiD method because it controls for trends that may affect both CJR and control group hospitals. In addition, we risk-adjusted estimates for beneficiary, market, and hospital characteristics that can vary over time and between the CJR and control group. The control group MSAs were weighted to be representative of the distribution of the mandatory CJR MSAs. The percent decrease in payments represents the percent change from the CJR baseline that is due to the CJR model. It is calculated by dividing the DiD estimate by the CJR baseline average.

---

[22] This hospital population excludes low-volume and rural hospitals in the mandatory CJR MSAs that were not required to continue in the model after PY2.

> **Standardized allowed amounts** include beneficiary cost sharing and do not include wage adjustments and other Medicare payment adjustments.

While the DiD method is intended to isolate the impact of the CJR model, it does not fully control for the unequal contribution of episodes from former Bundled Payments for Care Improvement (BPCI) participants in the CJR and control group samples. For hospitals that participated in BPCI, we excluded episodes when a hospital was participating in BPCI and included episodes after a hospital stopped participating in BPCI. Hospitals in CJR MSAs dropped out of BPCI earlier and at a higher rate than hospitals in control group MSAs. Hospitals in control group MSAs were less likely to drop out of BPCI, so more of their episodes are excluded from the CJR analysis. That led to a larger contribution of intervention episodes by former BPCI hospitals in the CJR group compared to the control group. Because former BPCI hospitals have lower average episode payments during the CJR performance period, their higher contribution to the CJR group may overstate the reduction in average CJR episode payments. As previously reported, our robustness analyses indicate that the larger contribution of episodes by former BPCI participants may overestimate the reductions in average episode payments due to CJR by roughly $200 per episode.[23]

We adjusted our methodology to account for two changes that affect our analysis beginning in the third performance year. First, CMS removed TKA from the inpatient only list in January 2018, allowing Medicare coverage for TKAs provided in the hospital outpatient setting. Second, the later part of the CJR intervention period (episodes starting in October 2018 or later) overlaps with the BPCI Advanced intervention period.

Accounting for the impact of the TKA policy change

Beginning in January 2018, Medicare removed TKA from the inpatient only list, which allowed coverage for the surgery when it was performed in the hospital outpatient setting. Following this policy change, both mandatory CJR and control group hospitals shifted some TKAs to the outpatient setting, however CJR hospitals shifted fewer. Although the share of TKAs performed in the hospital outpatient department continues to increase, the share of outpatient TKAs in mandatory CJR hospitals remains about 10 percentage points (pp) below the share in control group hospitals (Exhibit 3).

---

[23] Centers for Medicare & Medicaid Services. Comprehensive care for joint replacement model - third annual report. https://innovation.cms.gov/data-and-reports/2020/cjr-thirdannrpt. 2020:40

Exhibit 3:   Mandatory CJR hospitals have a lower proportion of outpatient TKAs than control group hospitals
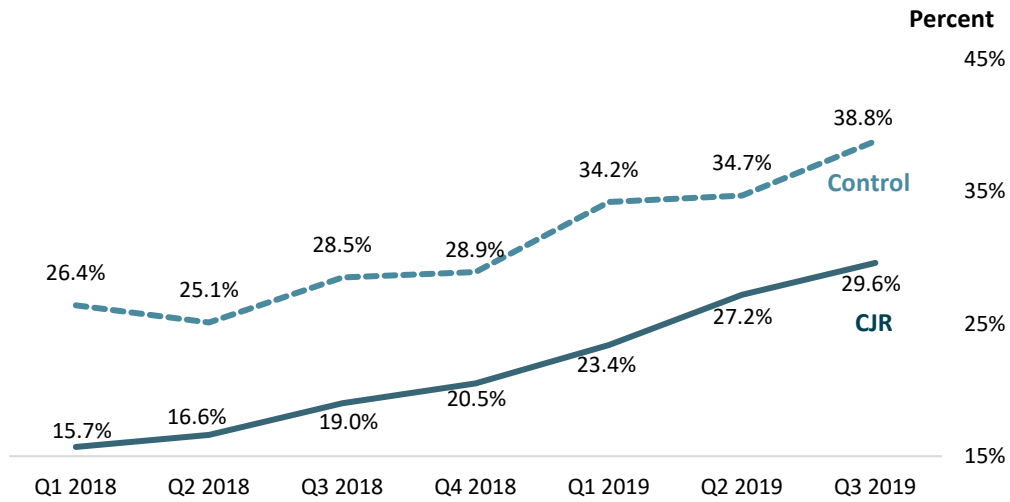
Note: TKA = total knee arthroplasty.

As a result of this differential response, the control group based on inpatient LEJR episodes is no longer an appropriate counterfactual for CJR episodes. Our analyses indicate that a portion of the CJR inpatient TKAs would have been outpatient in the absence of the CJR model (Appendix C, Section II.B.2 and Section IV). These episodes incurred higher episode payments as inpatient surgeries than they would have as outpatient surgeries for two reasons. First, the Medicare payment for the inpatient surgery is higher than the payment for the outpatient surgery.[24] Second, a fraction of the beneficiaries with an inpatient TKA were discharged to a SNF, but an outpatient surgery would not have triggered Medicare coverage of a SNF stay. Therefore, an appropriate counterfactual for the CJR episodes would include patients who would have received their TKA in the inpatient setting if they had been treated in a CJR hospital, but instead received their TKA in the outpatient setting at a control hospital.

For the third annual evaluation report, we provided two impact estimates: one based on inpatient episodes only, which is an overestimate of the true impact of the CJR model; and one based on inpatient episodes and outpatient TKAs (which include payments for the surgery and the following 90 days), which is an underestimate of the true impact of the CJR model on inpatient LEJRs.[25]

---

[24]   The Medicare 2018 base rate payment for MS-DRG 470—or the inpatient rate— was $12,380, whereas the payment for Ambulatory Payment Category 5115—or the outpatient rate— was $10,123.

[25]   Centers for Medicare & Medicaid Services. Comprehensive care for joint replacement model - third annual report. https://innovation.cms.gov/data-and-reports/2020/cjr-thirdannrpt. 2020:41

> ***An inpatient LEJR episode*** begins with an inpatient anchor hospitalization that meets CJR episode eligibility requirements and ends 90 days after discharge. Inpatient LEJR episodes are CJR episodes under the CJR model.
>
> ***Outpatient TKAs*** do not trigger CJR episodes under the CJR model. However, we constructed episodes that began with an outpatient TKA that extended for 90 days after the date of the surgery to use in our impact estimate.

For this fourth annual evaluation report, we have refined our methodology to provide a single estimate of the impact of the CJR model on LEJR episodes. We include all control outpatient TKAs and weight them to represent the proportion of CJR inpatient TKAs that are predicted to have been outpatient TKAs in the absence of the CJR model. This method provides an estimate of the impact of the CJR model on inpatient episodes only.

Accounting for the impact of BPCI Advanced contamination

The later part of the CJR intervention period (episodes starting in October 2018 or later) overlaps with the BPCI Advanced intervention period. Hospital participation in BPCI Advanced differs between the CJR and control groups because the CJR model takes precedence over the BPCI Advanced model. CJR hospitals are not eligible to participate in BPCI Advanced for LEJR and any episodes initiated by BPCI Advanced participating physician group practices (PGPs) at CJR hospitals are CJR episodes. In 2019, about 37% of control group episodes are BPCI Advanced episodes.

To control for BPCI Advanced contamination of the control group, we included an indicator variable for those episodes to account for the decline in the average payments of the control group in response to the BPCI Advanced model.

Additional details about the methodology are available in in Appendix C (Section III).

c. Results

During the first four performance years, the CJR model resulted in relative reductions in average episode payments (Appendix D). The analysis of average episode payments does not incorporate reconciliation payments made to CJR participant hospitals; therefore, the results do not represent savings to the Medicare program. An analysis of Medicare savings is presented in Section II.A.2.

LEJR episodes in mandatory CJR hospitals

Average payments for LEJR episodes declined for both mandatory CJR hospitals and control group hospitals during the first four performance years, although payments declined more for LEJR episodes initiated at mandatory CJR hospitals (Exhibit 4). Average payments decreased by $1,511 more for LEJR episodes initiated at mandatory CJR hospitals than for LEJR episodes initiated at control group hospitals from the baseline to the intervention period (p<0.01). This

relative reduction equates to a 5.2% decrease from the baseline in average payments for LEJR episodes at mandatory CJR hospitals.

This estimated reduction, based on including weighted outpatient TKAs in the control group, falls between the overestimated inpatient-only estimate ($1,716), and the underestimated all-inpatient-and-outpatient estimate ($1,407).

Exhibit 4:   Average payments declined more for LEJR episodes at mandatory CJR hospitals than for LEJR episodes at control group hospitals in PY1-4
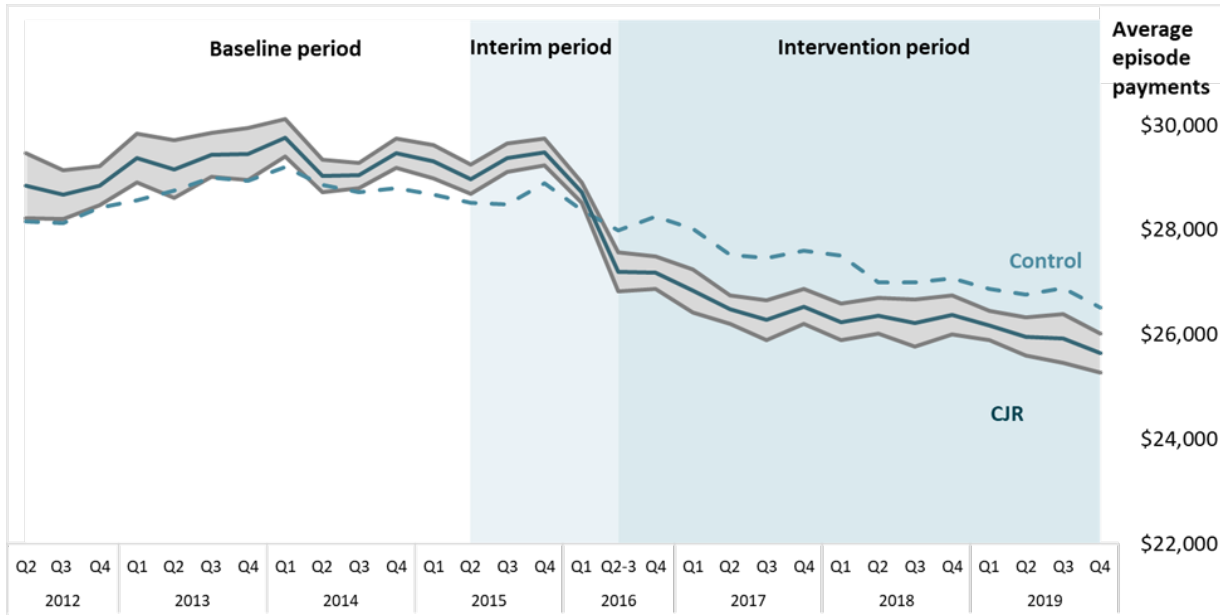
Source: CJR evaluation team analysis of Medicare claims and enrollment data for LEJR episodes initiated on or after January 2012 that ended by December 2019.

Notes: LEJR episodes that ended between April 1, 2015 and March 31, 2016 (the interim period) were excluded from our baseline because the CJR model was announced in July 2015 and hospitals may have been preparing for their future participation in the CJR model during that time.

Because CJR participant hospitals shifted a lower share of TKAs to the hospital outpatient setting, the control group includes outpatient TKA episodes that have been weighted to balance the episode volume in the CJR hospitals.

The gray shading represents the 90% confidence interval for the CJR estimate.

LEJR = lower extremity joint replacement, PY = performance year, TKA = total knee arthroplasty.

For mandatory CJR hospitals, average payments for LEJR episodes decreased in each of the four performance years. Average payments decreased by $1,431 more for LEJR episodes at mandatory CJR hospitals than at control group hospitals in the first performance year (p<0.01), by $1,618 more in the second performance year (p<0.01), by $1,330 more in the third performance year (p<0.01), and by $1,263 more in the fourth performance year (p<0.01, Exhibit 5).

Exhibit 5:   Average payments for LEJR episodes at mandatory CJR hospitals decreased in PY1, PY2, PY3, and PY4

Source: CJR evaluation team analysis of Medicare claims and enrollment data for LEJR episodes initiated in 2012 through 2014 that ended between April 2012 and March 2015 (baseline) and LEJR episodes initiated during or after April 2016 that ended by December 2019 (intervention).

Notes: The estimates in this exhibit are the result of a DiD model. DiD estimates that are significant at the 1%, 5%, or 10% significance level are indicated by red, orange, or yellow shaded circles, respectively. The whiskers represent 90% confidence intervals.
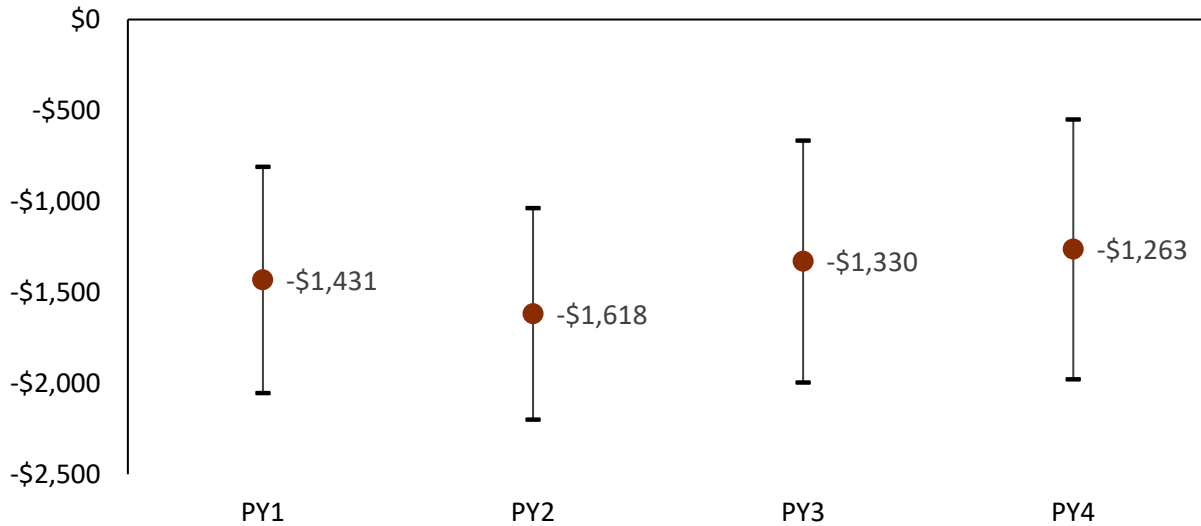
Because CJR participant hospitals shifted a lower share of TKAs to the hospital outpatient setting, the control group includes outpatient TKA episodes that have been weighted to balance the episode volume in the CJR hospitals.

DiD = difference-in-differences, LEJR = lower extremity joint replacement, PY = performance year, TKA = total knee arthroplasty.

d. Conclusion

Over the first four performance years, the CJR model resulted in relative reductions in average episode payments for mandatory CJR hospitals. The estimated reduction in payments is smaller in PY3 and PY4 than in PY2 because CJR participant hospitals shifted fewer TKAs to the outpatient setting than control hospitals. Outpatient episodes generally incur lower payments than inpatient episodes because the Medicare payment for the outpatient surgery is lower than the payment for inpatient surgery, and because outpatient surgeries are ineligible for Medicare coverage of a SNF stay. Thus, CJR payments were higher than they otherwise would have been because CJR hospitals retained a larger proportion of TKAs in the higher payment inpatient setting than the control group did.

On-going Medicare coverage changes and changes to the CJR episode definition will likely affect the impact of the CJR model. Medicare began covering outpatient TKAs in January 2018, and in January 2020, Medicare coverage expanded to include THAs in the hospital outpatient department. These coverage expansions may continue to dampen the impact of the CJR model if CJR hospitals retain more LEJRs in the inpatient setting than control hospitals. However, financial incentives for CJR hospitals will change beginning with the 3-year

extension of the CJR model in October 2021, when these outpatient LEJR episodes will be included in the model.[26] During the extension, the CJR model's site neutral target prices could affect the share of LEJRs in the hospital outpatient setting, which could affect future impact estimates. In addition, Medicare coverage expands to include TKAs in ambulatory surgery centers in January 2020 and THAs in ambulatory surgery centers in January 2021. LEJRs in ambulatory surgery centers will not be CJR episodes. These policy changes also will likely affect the impact of the CJR model.

2. *How much did the Medicare program save or lose due to the CJR model after accounting for reconciliation payments?*

Medicare achieves savings under the CJR model when reductions in episode payments are greater than the net reconciliation payments made to hospitals. This section presents estimated Medicare savings separately for mandatory CJR hospitals, opt-in CJR hospitals, and non-opt-in hospitals, as well as a combined estimate of Medicare savings for all hospitals that ever participated in the CJR model.

a. Key Findings

During the first four performance years, estimated Medicare savings for **all hospitals** (excluding low-volume or rural) during their participation in the CJR model was **$21.4 million,** ranging from losses of $75.0 million to savings of $117.9 million.

Mandatory CJR hospitals
- The CJR model generated estimated **savings of $76.0 million** (ranging from losses of $15.3 million to savings of $167.2 million)
- Estimated savings were **$494 per episode** (ranging from losses of $99 to savings of $1,087 per episode)

Opt-in CJR hospitals
- The CJR model generated estimated **losses of $44.5 million** (ranging from losses of $62.9 million to losses of $26.1 million)
- Estimated losses were **$843 per episode** (ranging from losses of $1,192 to $494 per episode)

Non-opt-in CJR hospitals
- The CJR model generated estimated **losses of $10.0 million** (ranging from losses of $20.2 million to savings of $0.2 million)
- Estimated losses were **$291 per episode** (ranging from losses of $590 to savings of $7 per episode)

---

[26]   Centers for Medicare & Medicaid Services. Medicare Program: Comprehensive Care for Joint Replacement Model Three-Year Extension and Changes to Episode Definition and Pricing; Medicare and Medicaid Programs; Policies and Regulatory Revisions in Response to the COVID–19 Public Health Emergency; Final Rule 2021:1-81.

b. Methods

As discussed in Section I, starting in PY3, the number of mandatory MSAs was scaled back to the 34 MSAs with the highest historical average episode payments. Our main analysis focuses on those MSAs and the 395 hospitals that were always required to participate (mandatory CJR hospitals). We also report results for hospitals in voluntary MSAs that opted to remain in the CJR model (opt-in CJR hospitals) for all four

Acronyms	
CI	confidence interval
DiD	difference-in-differences
LEJR	lower extremity joint replacement
MSA	metropolitan statistical area
PY	performance year
TKA	total knee arthroplasty

performance years and results for hospitals in voluntary MSAs that did not opt to remain in the CJR model (non-opt-in CJR hospitals) for the first two years in which they were participating. Our analysis excludes low volume and rural hospitals because CMS made their participation voluntary after the first two years of the model. As described in more detail in Section II.8.b, low-volume and rural hospitals differ in important ways from mandatory or voluntary CJR hospitals that would make them less comparable to our control groups. In Appendix E, we present our sensitivity analysis, which indicates that excluding low volume and rural hospitals from our Medicare program savings estimates is unlikely to affect our conclusions about the impact of the CJR model.

Medicare savings from the CJR model was calculated using the following formula:

Medicare savings = change in non-standardized paid amounts – reconciliation payments

The change in non-standardized paid amounts is estimated using our DiD methodology, which is explained fully in the Methods appendix (Appendix C, Section V.A).

To calculate Medicare savings, we used non-standardized paid amounts instead of the standardized allowed amounts used in the average episode and service-level payments impact analyses (Sections II.A.1 and II.A.3). Non-standardized paid amounts are the actual payments from Medicare to providers, so they incorporate geographic and other payment adjustments and exclude beneficiary cost-sharing, which aligns with how reconciliation payments are measured. The change in paid amounts reported here is different from the change in allowed amounts reported in the prior chapter. In general, the change in paid amounts is smaller because it does not include the change in beneficiary cost-sharing.

Reconciliation payments are the payments made to CJR participant hospitals by Medicare for meeting cost and quality targets and repayments from CJR participant hospitals to Medicare for failing to meet cost or quality targets.[27] See Appendix C, Section V for additional details.

---

[27] In the CJR model rule these payments are often referred to by the technical term "net payment reconciliation amounts" or "NPRA."

c. Results

Mandatory Hospitals

At mandatory hospitals, the CJR model reduced Medicare program spending by an estimated $76.0 million during the first four performance years, or about 1.82% of baseline payments for mandatory hospitals (Exhibit 6). Considering the uncertainty in the estimated reduction in per episode payments, savings due to the CJR model ranged from losses of $15.3 million to savings of $167.2 million. Medicare savings is based on an estimated reduction in per episode non-standardized paid amounts of $1,314 for the 153,813 LEJR episodes initiated at mandatory CJR hospitals, or $202.0 million.[28] Mandatory CJR hospitals received an average of $820 per episode in net reconciliation payments, or $126.1 million. Subtracting reconciliation payments from reductions in payments results in estimated savings of $494 per episode, ranging from losses of $99 to savings of $1,087.

Exhibit 6:   CJR likely resulted in Medicare savings from mandatory hospitals, PY1-4

Component	Estimate	Range (90% confidence)
Reduction in non-standardized paid amounts per episode	$1,314	$720 to $1,907
Reconciliation payments	$820	N/A
Medicare savings per episode	$494	-$99 to $1,807
Number of intervention episodes	153,813	N/A
Aggregate Medicare savings	$75,979,727	-$15,254,362 to $167,213,754

Source: CJR evaluation team analysis of Medicare claims and enrollment data for episodes initiated in 2012 through 2014 that ended between April 2012 and March 2015 (baseline) and episodes initiated during or after April 2016 that ended by December 2019 (intervention) and CJR payment contractor data for CJR participant hospitals in PY1-4.

Notes: Reductions in non-standardized paid amounts are based on a weighted average of performance year estimates from a DiD model of per-episode standardized paid amounts that have been multiplied by negative one and converted to non-standardized amounts. Ranges are based on 90% confidence intervals.

Because CJR participant hospitals shifted a lower share of TKAs to the hospital outpatient setting, the control group includes outpatient TKA episodes that have been weighted to balance the episode volume in the CJR hospitals.

DiD = difference-in-differences, PY = performance year, TKA = total knee arthroplasty.

We also evaluated Medicare savings separately for each performance year to better understand how model features that changed across performance years and how the removal of TKA from the inpatient-only list affected savings. The quality-adjusted target price decreased for 91% of mandatory hospitals in PY3 and 75% of mandatory hospitals in PY4. The quality-adjusted target prices changed because the CJR model is designed to shift from target prices based on historical hospital averages to regional averages and incorporate larger quality adjustments. Additionally, the stop-gain and stop-loss limits to reconciliation payments increased from 10% in PY3 to 20%

---

[28]   Excluding hospitals categorized as low-volume or rural, regardless of participant status.

in PY4. Finally, as noted above, Medicare began covering TKAs provided in a hospital outpatient setting in PY3.

The CJR model resulted in Medicare savings at mandatory CJR hospitals in PY2, and likely resulted in savings in PY1, PY3, and PY4. Medicare savings, however, decreased in PY3 and PY4 because mandatory CJR hospitals shifted fewer TKAs to the outpatient setting.[29] For mandatory CJR hospitals in PY1, Medicare savings is due to a reduction in per-episode non-standardized paid amounts of $1,394 less per-episode reconciliation payments of $872. This resulted in estimated Medicare savings of $522 per episode, ranging from losses of $56 to savings of $1,100 (Exhibit 7). For mandatory CJR hospitals in PY2, Medicare savings is due to a reduction in per-episode non-standardized paid amounts of $1,571 less per-episode reconciliation payments of $943. This resulted in estimated Medicare savings of $628 per episode, ranging from savings of $67 to $1,189. For mandatory CJR hospitals in PY3, Medicare savings is due to a reduction in per-episode non-standardized paid amounts of $1,254 less per-episode reconciliation payments of $698. This resulted in estimated Medicare savings of $556 per episode, ranging from losses of $101 to savings of $1,213. For mandatory CJR hospitals in PY4, Medicare savings is due to a reduction in per-episode non-standardized paid amounts of $1,098 less per-episode reconciliation payments of $785. This resulted in estimated Medicare savings of $313 per episode, ranging from losses of $397 to savings of $1,023.

Exhibit 7: The CJR model resulted in Medicare savings at mandatory hospitals in PY2, and likely resulted in savings in PY1, PY3, and PY4

Source: CJR evaluation team analysis of Medicare claims and enrollment data for episodes initiated in 2012 through 2014 that ended between April 2012 and March 2015 (baseline) and episodes initiated during or after April 2016 that ended by December 2019 (intervention) and CJR payment contractor data for CJR participant hospitals in PY1-4.

Notes: The arrows represent subtracting average reconciliation payments per episode from the reduction in paid amounts to calculate Medicare savings. Negative savings reflect Medicare losses.

The whiskers represent 90% confidence intervals. CIs are plotted as gray bars for reductions in non-standardized paid amounts and Medicare savings. Change in non-standardized paid amounts and CIs are based on estimates from a DiD model of per-episode standardized paid amounts that have been multiplied by negative one and converted to non-standardized amounts.

---

[29] As demonstrated in the third annual evaluation report, the reduction in average episode payments decreased when Medicare began covering TKAs in the hospital outpatient department. Centers for Medicare & Medicaid Services. Comprehensive care for joint replacement model - third annual report. https://innovation.cms.gov/data-and-reports/2020/cjr-thirdannrpt. 2020

Because CJR participant hospitals shifted a lower share of TKAs to the hospital outpatient setting, the control group includes outpatient TKA episodes that have been weighted to balance the episode volume in the CJR hospitals.

CI = confidence interval, DiD = difference-in-differences, PY = performance year, TKA = total knee arthroplasty.

Opt-in Hospitals

At opt-in hospitals in voluntary MSAs, the CJR model increased Medicare program spending by an estimated $44.5 million during the first four performance years, or approximately 3.87% of baseline payments for opt-in hospitals (Exhibit 8). Considering the uncertainty in the estimated reduction in per episode payments, losses ranged from $62.9 million to $26.1 million. These losses are based on an estimated reduction in per episode non-standardized paid amounts of $702 for the 52,813 LEJR episodes that were initiated at opt-in hospitals, or $37.1 million. From this, we subtracted average reconciliation payments of $1,545 per episode, or $81.6 million in total. This results in an estimated loss of $843 per episode, which ranges from a per episode loss of $1,192 to $494.

Exhibit 8:   The CJR model resulted in statistically significant Medicare losses at opt-in hospitals

Component	Estimate	Range (90% confidence)
Reduction in non-standardized paid amounts per episode	$702	$353 to $1,051
Reconciliation payments	$1,545	N/A
Medicare savings per episode	-$843	-$1,192 to -$494
Number of intervention episodes	52,813	N/A
Total Medicare savings	-$44,521,039	-$62,936,895 to -$26,105,178

Source: CJR evaluation team analysis of Medicare claims and enrollment data for episodes initiated in 2012 through 2014 that ended between April 2012 and March 2015 (baseline) and episodes initiated during or after April 2016 that ended by December 2019 (intervention) and CJR payment contractor data for CJR participant hospitals in PY1-4.

Notes: Reductions in non-standardized paid amounts are based on a weighted average of performance year estimates from a DiD model of per episode standardized paid amounts that have been multiplied by negative one and converted to non-standardized amounts. Ranges are based on 90% confidence intervals.

Because CJR participant hospitals shifted a lower share of TKAs to the hospital outpatient setting, the matched control group includes outpatient TKA episodes that have been weighted to balance the episode volume in the CJR hospitals.

DiD = difference-in-differences, PY = performance year, TKA = total knee arthroplasty.

The CJR model resulted in statistically significant Medicare losses at opt-in CJR hospitals in PY2, PY3, and PY4, and likely resulted in losses in PY1 (Exhibit 9). Medicare losses increased in PY3 and PY4 due to an increase in net reconciliation payments to opt-in hospitals in those years. For opt-in hospitals in PY1, Medicare losses are due to a reduction in per-episode non-standardized paid amounts of $516 less per-episode reconciliation payments of $950. This resulted in estimated Medicare losses of $435 per episode, ranging from losses of $925 to savings of $57. For opt-in hospitals in PY2, Medicare losses are due to a reduction in per-episode non-standardized paid amounts of $798 less per-episode reconciliation payments of $1,180. This resulted in estimated Medicare losses of $382 per episode, ranging from losses of $730 to $33. For opt-in hospitals in PY3, Medicare losses are due to a reduction in per-episode non-standardized paid amounts of $446 less per-episode reconciliation payments of $1,467.

This resulted in estimated Medicare losses of $1,021 per episode, ranging from losses of $1,437 to $605. For opt-in hospitals in PY4, Medicare losses are due to a reduction in per-episode non-standardized paid amounts of $953 less per-episode reconciliation payments of $2,389. This resulted in estimated Medicare losses of $1,436 per episode, ranging from losses of $1,922 to $949.

Exhibit 9:   The CJR model likely resulted in Medicare losses at opt-in hospitals in PY1, and resulted in losses in PY2, PY3, and PY4

Source: CJR evaluation team analysis of Medicare claims and enrollment data for episodes initiated in 2012 through 2014 that ended between April 2012 and March 2015 (baseline) and LEJRs initiated during or after April 2016 that ended by December 2019 (intervention) and CJR payment contractor data for CJR participant hospitals in PY1-4.

Notes: The arrows represent subtracting average reconciliation payments per episode from the reduction in paid amounts to calculate Medicare savings. Negative savings reflect Medicare losses.

The whiskers represent 90% confidence intervals. CIs are plotted as gray bars for reductions in non-standardized paid amounts and Medicare savings. Change in non-standardized paid amounts and CIs are based on estimates from a DiD model of per episode standardized paid amounts that have been multiplied by negative one and converted to non-standardized amounts.
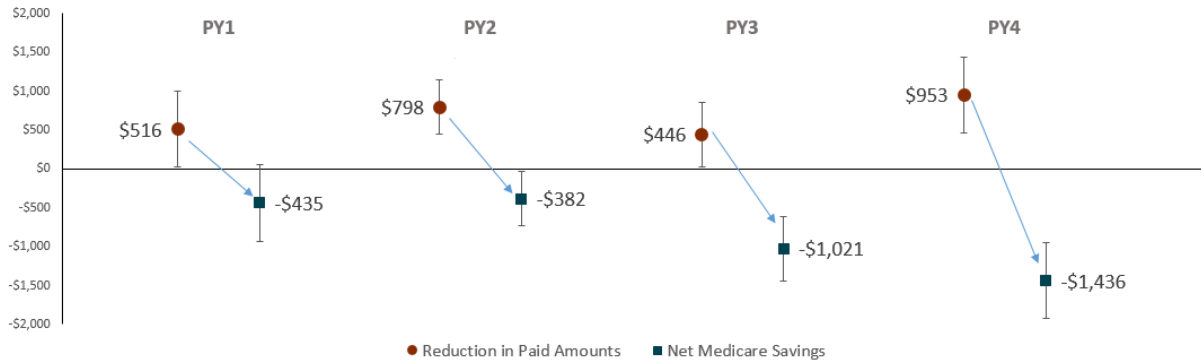
Because CJR participant hospitals shifted a lower share of TKAs to the hospital outpatient setting, the control group includes outpatient TKA episodes that have been weighted to balance the episode volume in the CJR hospitals.

CI = confidence interval, DiD = difference-in-differences, LEJR = lower extremity joint replacement, PY = performance year, TKA = total knee arthroplasty.

Non-opt-in Hospitals

At non-opt-in hospitals in voluntary MSAs, the CJR model increased Medicare program spending by an estimated $10.0 million during the first *two* performance years, or approximately 1.25% of baseline payments for non-opt-in hospitals (Exhibit 10). We consider only the first two performance years because these hospitals were no longer in the CJR model starting in PY3. The losses are based on an estimated reduction in per episode non-standardized paid amounts of $370 for the 34,312 LEJR episodes initiated at non-opt-in hospitals during the first two performance years, or $12.7 million. Non-opt-in hospitals received an average of $661 per episode in reconciliation payments during PY1 and PY2, or $22.7 million in total. Subtracting reconciliation payments from reductions in payments results in an estimated loss of $291 per episode, ranging from losses of $590 to savings of $7 per episode.

Exhibit 10: The CJR model likely resulted in losses to Medicare at non-opt-in hospitals

Component	Estimate	Range (90% confidence)
Reduction in non-standardized paid amounts per episode	$370	$71 to $669
Reconciliation payments	$661	N/A
Medicare savings per episode	-$291	-$590 to $7
Number of intervention episodes	34,312	N/A
Total Medicare savings	-$9,999,202	-$20,242,085 to $243,678

Source: CJR evaluation team analysis of Medicare claims and enrollment data for episodes initiated in 2012 through 2014 that ended between April 2012 and March 2015 (baseline) and episodes initiated during or after April 2016 that ended by December 2019 (intervention) and CJR payment contractor data for CJR participant hospitals in PY1-4.

Notes: Reductions in non-standardized paid amounts are based on a weighted average of performance year estimates from a DiD model of per episode standardized paid amounts that have been multiplied by negative one and converted to non-standardized amounts. Ranges are based on 90% confidence intervals.

Because CJR participant hospitals shifted a lower share of TKAs to the hospital outpatient setting, the control group includes outpatient TKA episodes that have been weighted to balance the episode volume in the CJR hospitals.

DiD = difference-in-differences, PY = performance year, TKA = total knee arthroplasty.

The CJR model resulted in Medicare losses at non-opt-in hospitals in PY1, and likely resulted in losses in PY2 (Exhibit 11). We do not report estimated Medicare savings for non-opt-in hospitals in PY3 and PY4 because these hospitals were no longer part of the model in those years. For non-opt-in hospitals in PY1, Medicare losses are due to a reduction in per-episode non-standardized paid amounts of $74, less per-episode reconciliation payments of $497. This resulted in estimated Medicare losses of $423, ranging from losses of $788 to $58. For non-opt-in hospitals in PY2, Medicare losses are due to a reduction in per-episode non-standardized paid amounts of $511, less per-episode reconciliation payments of $739. This resulted in estimated Medicare losses of $229, ranging from losses of $532 to savings of $75.

Exhibit 11: The CJR model resulted in Medicare losses in PY1 and likely resulted in Medicare losses in PY2 for non-opt-in hospitals

Source: CJR evaluation team analysis of Medicare claims and enrollment data for episodes initiated in 2012 through 2014 that ended between April 2012 and March 2015 (baseline) and LEJRs initiated during or after April 2016 that ended by December 2019 (intervention) and CJR payment contractor data for CJR participant hospitals in PY1-4.
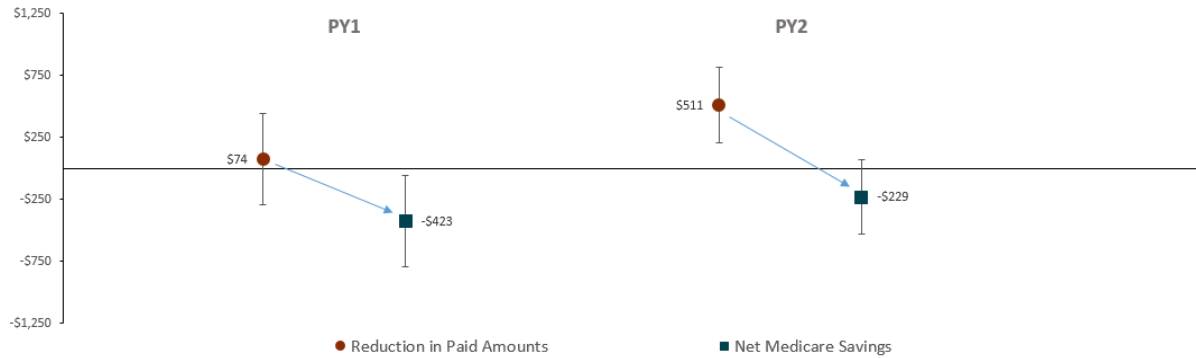
Notes: The arrows represent subtracting average reconciliation payments per episode from the reduction in paid amounts to calculate Medicare savings. Negative savings reflect Medicare losses.

The whiskers represent 90% confidence intervals. CIs are plotted as gray bars for reductions in non-standardized paid amounts and Medicare savings. Change in non-standardized paid amounts and CIs are based on estimates from a DiD model of per episode standardized paid amounts that have been multiplied by negative one and converted to non-standardized amounts.

CI = confidence interval, DiD = difference-in-differences, LEJR = lower extremity joint replacement, PY = performance year.

All Hospitals

Across all hospitals that ever participated, our estimate suggests that the CJR model reduced Medicare program spending by $21.4 million during the first four performance years, or approximately 0.35% of baseline payments (Exhibit 12).[30] Considering the uncertainty in the estimated reduction in per-episode payments, savings due to the CJR model ranged from losses of $75.0 million to savings of $117.9 million. Due to this wide range, which includes the possibility of substantial losses or savings, we cannot conclude that the CJR model resulted in savings to Medicare across all hospitals that ever participated. Medicare savings is based on an estimated reduction in per-episode non-standardized paid amounts of $1,045 for the 240,938 LEJR episodes initiated at all CJR hospitals, or $251.8 million. CJR hospitals received an average of $956 per episode in net reconciliation payments, or $230.4 million. Subtracting reconciliation payments from reductions in payments results in estimated savings of $89 per episode, ranging from losses of $311 to savings of $489.

---

[30]   This analysis excludes low volume and rural hospitals, regardless of participant status. In Appendix E, we report results from a sensitivity analysis which concludes that inclusion of low volume and rural participant hospitals would not significantly change our main conclusions.

Exhibit 12: We cannot conclude that the CJR model resulted in savings to Medicare across all hospitals and all performance years

Component	Estimate	Range (90% confidence)
Reduction in non-standardized paid amounts per episode	$1,045	$645 to $1,445
Reconciliation payments	$956	N/A
Medicare savings per episode	$89	-$311 to $489
Number of intervention episodes	240,938	N/A
Total Medicare savings	$21,437,936	-$75,033,091 to $117,908,963

Source: CJR evaluation team analysis of Medicare claims and enrollment data for episodes initiated in 2012 through 2014 that ended between April 2012 and March 2015 (baseline) and episodes initiated during or after April 2016 that ended by December 2019 (intervention) and CJR payment contractor data for CJR participant hospitals in PY1-4.

Notes: Reductions in non-standardized paid amounts are based on a weighted average of performance year estimates from a DiD model of per-episode standardized paid amounts that have been multiplied by negative one and converted to non-standardized amounts. Ranges are based on 90% confidence intervals.

Because mandatory and opt-in CJR hospitals shifted a lower share of TKAs to the hospital outpatient setting, the respective control groups include outpatient TKA episodes that have been weighted to balance the episode volume in the CJR hospitals. The control group for opt-in hospitals is constructed through a matching process. The control group for non-opt-in hospitals includes only inpatient TKAs.

DiD = difference-in-differences, PY = performance year, TKA = total knee arthroplasty.

We cannot conclude that the CJR model resulted in Medicare savings across all hospitals in any performance year (Exhibit 13). While our point estimates indicate there may have been savings in some years, the ranges around the estimates are wide and include the possibility of losses or savings, which precludes us from concluding there were savings to Medicare. Medicare savings decreased in PY4 due to increased net reconciliation payments and due to control hospitals performing more TKAs in the outpatient setting (which lead to lower payments in the control group). For all hospitals in PY1, Medicare savings are due to a reduction in per-episode non-standardized paid amounts of $862 less per-episode reconciliation payments of $782. This resulted in estimated Medicare savings of $80 per episode, ranging from losses of $277 to savings of $437. For all hospitals in PY2, Medicare savings are due to a reduction in per-episode non-standardized paid amounts of $1,122 less per-episode reconciliation payments of $934. This resulted in estimated Medicare savings of $187 per episode, ranging from losses of $148 to savings of $523. For all hospitals in PY3, Medicare savings are due to a reduction in per-episode non-standardized paid amounts of $1,037 less per-episode reconciliation payments of $905. This resulted in estimated Medicare savings of $132, ranging from losses of $374 to savings of $638. For all hospitals in PY4, Medicare savings are due to a reduction in per-episode non-standardized paid amounts of $1,066 less per-episode reconciliation payments of $1,138. This resulted in estimated Medicare losses of $72 per episode, ranging from losses of $636 to savings of $492.

Exhibit 13: Across all hospitals, we cannot conclude that the CJR model resulted in net savings in any performance year

Source: CJR evaluation team analysis of Medicare claims and enrollment data for episodes initiated in 2012 through 2014 that ended between April 2012 and March 2015 (baseline) and episodes initiated during or after April 2016 that ended by December 2019 (intervention) and CJR payment contractor data for CJR participant hospitals in PY1-4.
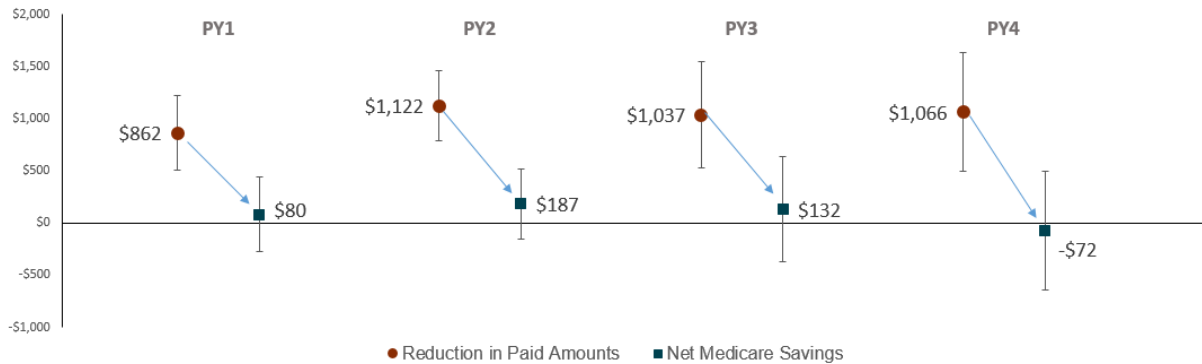
Notes: The arrows represent subtracting average reconciliation payments per episode from the reduction in paid amounts to calculate Medicare savings. Negative savings reflect Medicare losses.

The whiskers represent 90% confidence intervals. CIs are plotted as gray bars for reductions in non-standardized paid amounts and Medicare savings. Change in non-standardized paid amounts and CIs are based on estimates from a DiD model of per-episode standardized paid amounts that have been multiplied by negative one and converted to non-standardized amounts.

Because mandatory and opt-in CJR hospitals shifted a lower share of TKAs to the hospital outpatient setting, the respective control groups include outpatient TKA episodes that have been weighted to balance the episode volume in the CJR hospitals. The control group for opt-in hospitals is constructed through a matching process. The control group for non-opt-in hospitals includes only inpatient TKAs.

CI = confidence interval, DiD = difference-in-differences, PY = performance year, TKA = total knee arthroplasty.

d.   Conclusion

Over the entire CJR model period, considering all hospitals that ever participated, excluding hospitals designated as low volume or rural, estimated total Medicare savings was $21.4 million, with an estimated range of $75.0 million in losses to $117.9 million in savings. Given this wide range, which includes the possibility of substantial losses or savings, we cannot conclude that the CJR model resulted in savings overall.

Medicare savings estimates due to the CJR model vary across the different participation status of hospitals across the first four performance years. Hospitals that were required to participate for the entire period, or mandatory hospitals, reduced CJR episode payments by an estimated $202.0 million, which ranged from $110.8 million to $293.3 million. They received $126.1 million in net reconciliation payments. As a result, hospitals that were required to participate in the CJR model for the first four performance years achieved estimated net Medicare savings of $76.0 million, which ranged from losses of $15.3 million to savings of $167.2 million.

Opt-in hospitals in voluntary MSAs, who chose to remain in the CJR model for the latter two performance years, contributed to Medicare losses. At the start of the model, about 73% of these

hospitals had average historical payments below their PY1 target price.[31] These hospitals needed to make no or little change to their average payments in to receive reconciliation payments. Opt-in hospitals in voluntary MSAs reduced CJR episode payments by an estimated $37.1 million, which ranged from $55.5 million to $18.7 million. They received $81.6 million in net reconciliation payments. As a result, hospitals in voluntary MSAs that opted to remain in the CJR model caused statistically significant estimated net Medicare losses of $44.5 million, which ranged from losses of $62.9 million to losses of $26.1 million.

Non-opt-in hospitals in voluntary MSAs, which only participated during the first two performance years, likely also contributed to Medicare losses, although they had the lowest average reconciliation payments of the three hospital groups. Non-opt-in hospitals in voluntary MSAs achieved an estimated reduction in CJR episode payments of $12.7 million, ranging from $2.5 million to $22.9 million. They received $22.7 million in net reconciliation payments, resulting in estimated net Medicare losses of $10.0 million, ranging from losses of $20.2 million to savings of $0.2 million.

3. *What was the impact of the CJR model on service-level payments and service use?*

Changes in service-level payments and use provide insights into how hospitals reduced average episode payments. Payments for PAC, which comprise roughly one-third of LEJR episode payments, can be reduced by shifting service use from more to less intensive care settings that receive lower Medicare payments. Generally, average IRF payments are higher than average SNF payments, and Medicare payments for both of these institutional PAC settings tend to be higher than payments for home health (HH) care. Prior to the outpatient TKA coverage change, hospital payments were unlikely to change because hospitals receive a per-discharge payment for inpatient LEJRs that typically is not affected by length of stay or services provided during the hospitalization. The Medicare payment for a TKA performed in the outpatient setting, however, is lower than the inpatient payment. Because the CJR model affected the distribution of inpatient versus outpatient procedures, hospital payments could change in response to the CJR model.

a. *Key Findings*

- The relative decrease in average episode payments was driven by relative decreases in IRF and SNF payments.

- A smaller proportion of CJR patients than control patients were first discharged to an IRF or SNF and a larger proportion were first discharged to an HHA.

---

[31]   Centers for Medicare & Medicaid Services. Comprehensive care for joint replacement model - third annual report. https://innovation.cms.gov/data-and-reports/2020/cjr-thirdannrpt. 2020

b. Methods

This analysis uses a DiD design (described in Section II.A.1.b) to estimate the differential change in average standardized allowed amounts (payments) and average utilization by service during the 90 days following discharge from the hospital for mandatory CJR hospitals.[32,33] Average payments by service are calculated across all episodes, including episodes that did not receive the particular service, while average length of stay and number of visits are based only on the episodes that used that particular service.

Acronyms	
DiD	difference-in-differences
HH	home health
HHA	home health agency
IPPS	inpatient prospective payment system
IRF	inpatient rehabilitation facility
LEJR	lower extremity joint replacement
LOS	length of stay
PAC	post-acute care
PDGM	Patient-Driven Groupings Model
PDPM	Patient-Driven Payment Model
pp	percentage point
PY	performance year
SNF	skilled nursing facility
TKA	total knee arthroplasty

c. Results

During the first four performance years, the relative decrease in average episode payments ($1,511, $p<0.01$) for mandatory CJR hospitals was primarily driven by relative reductions in IRF and SNF payments. There were also relative reductions in readmission payments. Relative to control group episodes from the baseline to the intervention period, average IRF payments for CJR episodes decreased by $593 ($p<0.01$), or 26.5% from the CJR baseline, SNF payments decreased by $843 ($p<0.01$), or 13.7% from the CJR baseline, and readmission payments decreased by $153 ($p<0.05$), or 12.5% from the CJR baseline (Exhibit 14). While not statistically significant at the 10% level, there were also relative increases in anchor payments ($p=0.13$, Appendix D). We expect relative increases in anchor payments to continue to grow as the shift to outpatient TKA continues because of the lower anchor payments in the outpatient setting.

---

[32] Appendix F contains additional detail about how outcomes are defined.

[33] We used standardized payments to ensure that observed payment differences reflect actual differences in billed services rather than Medicare payment policies. We used allowed amounts to eliminate variation in payments due to whether beneficiaries have met their deductible when they had the LEJR surgery. The change in standardized allowed amounts reported here is different from the change in non-standardized paid amounts reported in the prior chapter. In general, the change in standardized allowed amounts is greater because it includes the change in beneficiary cost-sharing.

Exhibit 14: The reduction in average episode payments was driven by decreases in inpatient rehabilitation facility and skilled nursing facility payments, PY1-4

Relative change in average payments

Source: CJR evaluation team analysis of Medicare claims and enrollment data for episodes initiated in 2012 through 2014 that ended between April 2012 and March 2015 (baseline) and episodes initiated during or after April 2016 that ended by December 2019 (intervention).

Notes: The estimates in this exhibit are the result of a DiD model. DiD estimates that are significant at the 1%, 5%, or 10% significance level are indicated by red, orange, or yellow shaded bars, respectively.

Because CJR participant hospitals shifted a lower share of TKAs to the hospital outpatient setting, the control group includes outpatient TKA episodes that have been weighted to balance the episode volume in the CJR hospitals.

DiD = difference-in-differences, HHA = home health agency, IRF = inpatient rehabilitation facility, PY = performance year, SNF = skilled nursing facility, TKA = total knee arthroplasty.

⚠ We cannot be certain that there is no impact of the model because this outcome failed parallel trends tests (Appendix K). Parallel trends is an assumption that underlies our methodological approach, and without it we do not necessarily believe the control group acts as an accurate representation of what would have occurred in CJR hospitals in the absence of the CJR model. Please see Appendix C (Section III.C.1.c) for additional details.

The relative decrease in IRF payments is the result of a relative reduction in the proportion of CJR patients discharged from the hospital to an IRF. Although the proportion of patients discharged to an IRF declined for both groups, the decline was 3.9pp greater for CJR patients than for control patients, representing a 28.1% decrease from the CJR baseline proportion (p<0.01, Exhibit 15). Among patients with an IRF stay, there was no relative change in the number of days that CJR patients spent in an IRF (Appendix D). However, this is not unexpected because reducing length of stay (LOS) does not affect Medicare's per discharge IRF payment.

**Exhibit 15: The reasons for the decrease in post-acute
care payments differ by setting**

Setting	Relative Change in		
	Payment	Proportion of patients first discharged	Number of days/visits
Inpatient rehabilitation facility	-$593	-28.1%	-0.1
Skilled nursing facility	-$843	-6.5%	-2.6
Home health agency	+$65 ⚠	+20.5%	-1.0

Source: CJR evaluation team analysis of Medicare claims and enrollment
data for episodes initiated in 2012 through 2014 that ended between
April 2012 and March 2015 (baseline) and episodes initiated during
or after April 2016 that ended by December 2019 (intervention).

Notes: The estimated relative change in utilization is the result of a DiD
model. DiD estimates that are significant at the 1%, 5%, or 10%
significance level are indicated by red, orange, or yellow shaded
shapes, respectively. All other estimates are indicated by gray
shapes. Hashed shapes reflect measures of utilization that do not
influence payment.

The change in the proportion of patients first discharged to each
PAC setting represents the percent change from the CJR baseline
that is due to CJR. It is calculated by dividing the DiD estimate by
the CJR baseline average.

For SNF LOS, beneficiaries must have spent at least one day in the
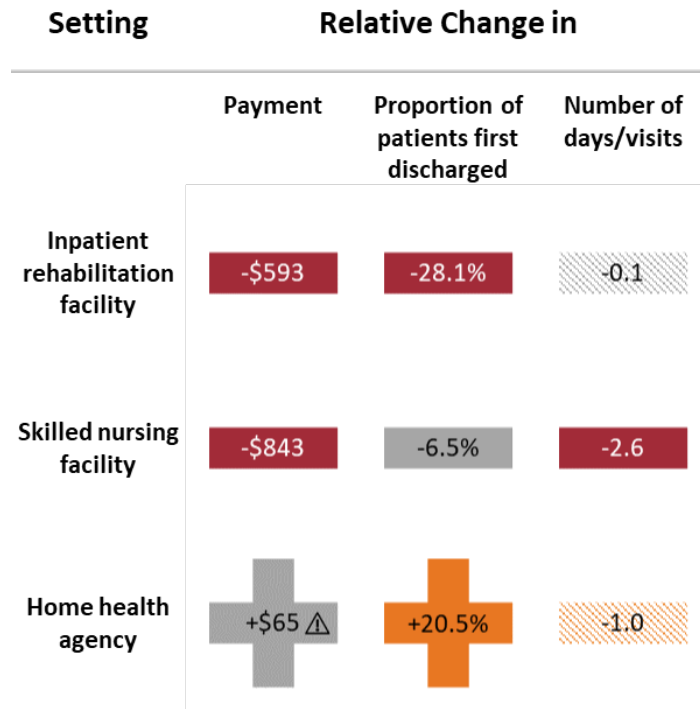SNF setting.

Because CJR participant hospitals shifted a lower share of TKAs to
the hospital outpatient setting, the control group includes
outpatient TKA episodes that have been weighted to balance the
episode volume in the CJR hospitals.

DiD = difference-in-differences, LOS = length of stay, PAC = post-
acute care, SNF = skilled nursing facility, TKA = total knee
arthroplasty.

⚠ We cannot be certain that there is no impact of the model because
this outcome failed parallel trends tests (Appendix K). Parallel trends
is an assumption that underlies our methodological approach, and
without it we do not necessarily believe the control group acts as an
accurate representation of what would have occurred in CJR
hospitals in the absence of the CJR model. Please see Appendix C
(Section III.C.1.c) for additional details.

The relative decrease in SNF payments is due to a relative reduction in the proportion of CJR patients discharged from the hospital to a SNF and a relative decrease in the number of SNF days among patients with a SNF stay. Medicare pays SNFs per diem amounts, so reducing SNF days results in lower episode payments. The proportion of patients first discharged to a SNF decreased by 2.7pp more for CJR episodes than for control group episodes from the baseline to the intervention period, representing a 6.5% decrease from the CJR baseline (p=0.103, Exhibit 15 and 17). While this estimate does not fall below the typical 10% significance level, we believe the estimate captures a true effect of the CJR model because it is consistent with previous, statistically significant estimates.[34] For patients with a SNF stay, the average length of stay decreased by 2.6 days more for CJR than for control group episodes from the baseline to the intervention period, representing a 9.5% decrease from the CJR baseline (p<0.01, Exhibit 15).

The proportion of patients first discharged to an HHA increased by 7.5pp more for CJR episodes than for control group episodes, representing a 20.5% change from the CJR baseline proportion (p<0.05, Exhibits 15 and 18). We cannot be certain that there were no impacts of the model on HHA payments (Exhibit 15) and the proportion of patients who receive any HH care (Appendix D) because both outcomes failed parallel trends tests. Parallel trends is an assumption that underlies our methodological approach. Because failure to pass the parallel trends test indicates that CJR and control group hospitals may have had different patterns of HHA payments and HH use in the baseline, we cannot isolate the impact of the CJR model on these outcomes. Please see Appendix C (Section III.C.1.c) for additional details.

---

[34]   We believe the change in statistical significance is likely due to the differential response to the outpatient TKA policy because the estimate is larger and statistically significant at the 5% level when the sample contains only inpatient episodes.

Exhibit 16: The proportion of patients discharged to inpatient rehabilitation facilities declined more for CJR than for control episodes, PY1-4

Source: CJR evaluation team analysis of Medicare claims and enrollment data for episodes initiated on or after January 2012 that ended by December 2019.

Notes: Episodes that were initiated in calendar year 2015 and ended between April 1, 2015 and March 31, 2016 (the interim period) were excluded from our baseline because the CJR model was announced in July 2015 and hospitals were likely preparing for their future participation in the CJR model during that time.

Because CJR participant hospitals shifted a lower share of TKAs to the hospital outpatient setting, the control group includes outpatient TKA episodes that have been weighted to balance the episode volume in the CJR hospitals.

The gray shading represents the 90% confidence interval for the CJR estimate.

PY = performance year, TKA = total knee arthroplasty.

Exhibit 17: The proportion of patients discharged to skilled nursing facilities declined more for CJR than for control episodes, PY1-4

Source: CJR evaluation team analysis of Medicare claims and enrollment data for episodes initiated on or after January 2012 that ended by December 2019.

Notes: Episodes that were initiated in calendar year 2015 and ended between April 1, 2015 and March 31, 2016 (the interim period) were excluded from our baseline because the CJR model was announced in July 2015 and hospitals were likely preparing for their future participation in the CJR model during that time.

Because CJR participant hospitals shifted a lower share of TKAs to the hospital outpatient setting, the control group includes outpatient TKA episodes that have been weighted to balance the episode volume in the CJR hospitals.

The gray shading represents the 90% confidence interval for the CJR estimate.

PY = performance year, TKA = total knee arthroplasty.

Exhibit 18: The proportion of patients first discharged to home health agencies increased more for CJR episodes than for control episodes, PY1-4

Source: CJR evaluation team analysis of Medicare claims and enrollment data for episodes initiated on or after January 2012 that ended by December 2019.
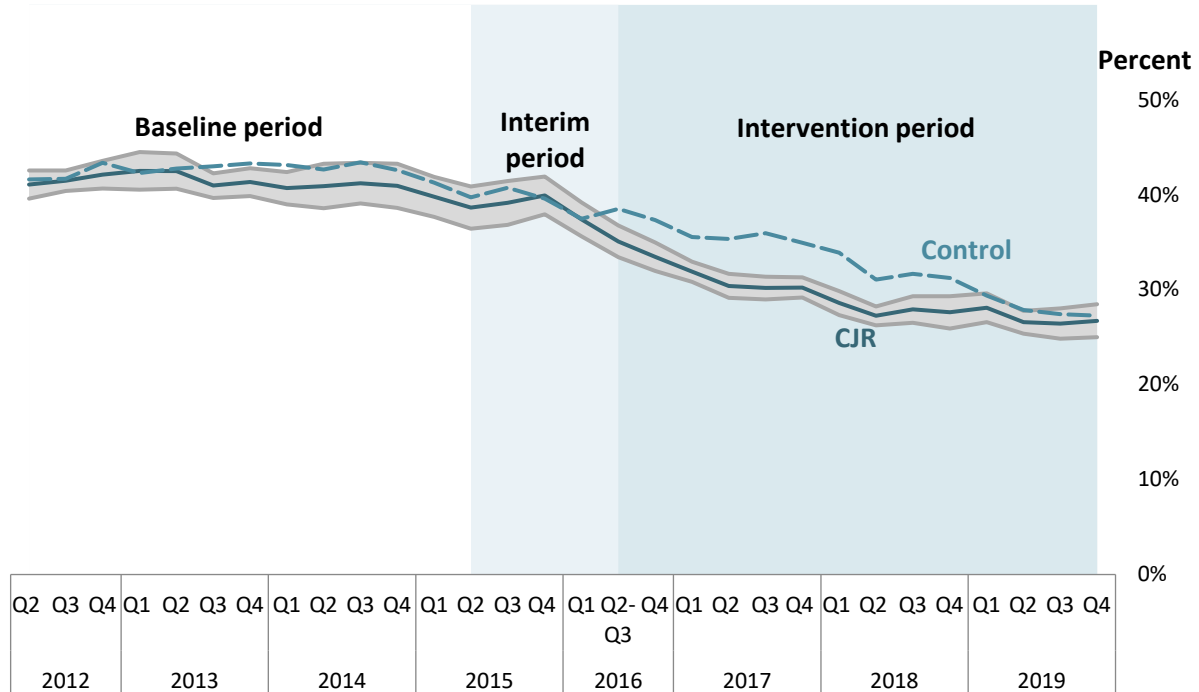
Notes: Episodes that were initiated in calendar year 2015 and ended between April 1, 2015 and March 31, 2016 (the interim period) were excluded from our baseline because the CJR model was announced in July 2015 and hospitals were likely preparing for their future participation in the CJR model during that time.

Because CJR participant hospitals shifted a lower share of TKAs to the hospital outpatient setting, the control group includes outpatient TKA episodes that have been weighted to balance the episode volume in the CJR hospitals.

The gray shading represents the 90% confidence interval for the CJR estimate.

PY = performance year, TKA = total knee arthroplasty.

The relative decrease in the proportion of CJR patients first discharged to institutional PAC and the increase in the proportion first discharged to an HHA suggest hospitals responded to the CJR model by shifting beneficiaries away from more intensive PAC settings to less intensive ones.

The relative decrease in average readmission payments was primarily due to changes made by a few CJR participant hospitals, rather than a reduction in the overall readmission rate. During the baseline, relatively more CJR patients were discharged to IPPS hospitals for rehabilitation. In the intervention period, this practice mostly stopped, resulting in a relative decrease in readmission payments that was driven by changes made by a few CJR participant hospitals, rather than more widespread decreases in readmission rates. Specifically, 64 CJR participant hospitals in the baseline discharged 1,886 LEJR patients, or 1.3% of baseline episodes, to IPPS hospitals under MS-DRG 945 or 946 (rehabilitation with or without complications or comorbidities). The majority (94%) of these patients were discharged to six hospitals for rehabilitation. In the control group, 20 hospitals discharged 0.02% of LEJR patients to IPPS hospitals for rehabilitation during the baseline.

d. Conclusion

CJR participant hospitals decreased LEJR episode payments by reducing the use of more intensive institutional PAC services. The proportion of LEJR patients who received IRF and SNF care decreased, and for those who received SNF care, the number of days in the SNF went down. The relative proportion of patients who were discharged from the hospital to an HHA went up.

SNF and HHA Medicare Payment Changes

CMS made changes to the SNF and HHA payment systems effective as of October 1, 2019, which will likely affect SNF and HHA utilization and payments. The SNF Patient-Driven Payment Model (PDPM) and the Home Health Patient-Driven Groupings Model (PDGM) both use clinical and functional characteristics of patients rather than volume of therapy services furnished to classify patients for payment, which reduces incentives to provide more therapy to boost payments. In addition, as of October 1, 2018, Medicare SNF rates were adjusted based on their 30-day All-Cause Readmission Measure, bringing incentives into alignment with CJR hospitals that must meet specified quality measure performance targets to qualify for realized savings. In the future, we will investigate whether and how these payment reforms affect CJR participant hospitals differently than non-CJR hospitals.

---

4. What was the impact of the CJR model on quality of care?

The CJR model was designed to reward hospitals that delivered high quality of care by applying a lower discount to the target price for participant hospitals with a higher quality score, which results in a higher quality-adjusted target price.

---

a. Key Findings

- The unplanned readmission rate decreased under the CJR model, while the emergency department use rate and the mortality rate remained unchanged.

- The complication rate decreased for elective LEJR episodes under the CJR model.

---

b. Methods

This analysis used the DiD approach described in Section II.A.1.b to estimate the relative change in outcomes for mandatory CJR hospitals. Appendix C (Section III) includes more detailed information about the methodology.

Acronyms

DiD	difference-in-differences
ED	emergency department
LEJR	lower extremity joint replacement
pp	percentage point
PY	performance year
TKA	total knee arthroplasty

c. Results

During the first four performance years, for mandatory CJR hospitals, the CJR model resulted in relative reductions in the 90-day unplanned readmission rate. There was no change in the 90-day ED use rate or the mortality rate during the anchor hospitalization plus the 90-day post-discharge period (Exhibit 19).

The unplanned readmission rate decreased more for CJR episodes than for control group episodes, representing a 3.5% decrease from the CJR baseline (Exhibit 19, p<0.10).

Neither the change in ED use nor the change in the mortality rate were statistically different between mandatory CJR and control hospitals, however, ED use and the mortality rate for CJR episodes increased by 1.0% from the CJR baseline. Because of the importance of these quality measures, we further examined the estimated changes. For mandatory CJR hospitals, ED use increased from 13.1% in the baseline to 14.1% in the intervention period. For control hospitals, it increased from 12.7% to 13.6% during the same time (Appendix D, Exhibit D-1). During the intervention period both mandatory CJR hospitals and control hospitals increased ED use, with mandatory CJR hospitals increasing by 0.1pp more than control hospitals. Though this difference is small and not statistically significant, we will continue to monitor ED use. For mandatory CJR hospitals, the mortality rate decreased from 2.6% in the baseline to 2.5% in the intervention period. For control hospitals, it decreased from 2.7% to 2.5% during the same period (Appendix D, Exhibit D-1). The control group had a higher baseline mortality rate, but decreased the mortality rate to approximately the same level as that of the CJR group during the intervention period.

We also evaluated changes in the complication rate, which is specific to elective LEJRs.[35] The complication rate for elective LEJR decreased more for CJR episodes than control group episodes, representing a 7.9% decrease from the CJR baseline (Exhibit 19, p<0.05). The complication rate measures the proportion of elective episodes with a complication during the anchor hospitalization or a readmission, so the reduction in the complication rate is consistent with the reduction in readmissions.

These results align with qualitative findings from prior reports on hospital efforts to reduce readmission rates and complication rates through improved care coordination.[36]

---

[35] The complication rate applies to all elective LEJRs, including inpatient and outpatient LEJRs.

[36] Centers for Medicare & Medicaid Services. Comprehensive care for joint replacement model - third annual report. https://innovation.cms.gov/data-and-reports/2020/cjr-thirdannrpt. 2020:122-127

Exhibit 19: Quality of care was maintained or improved under the CJR model for mandatory CJR hospitals, PY1-4
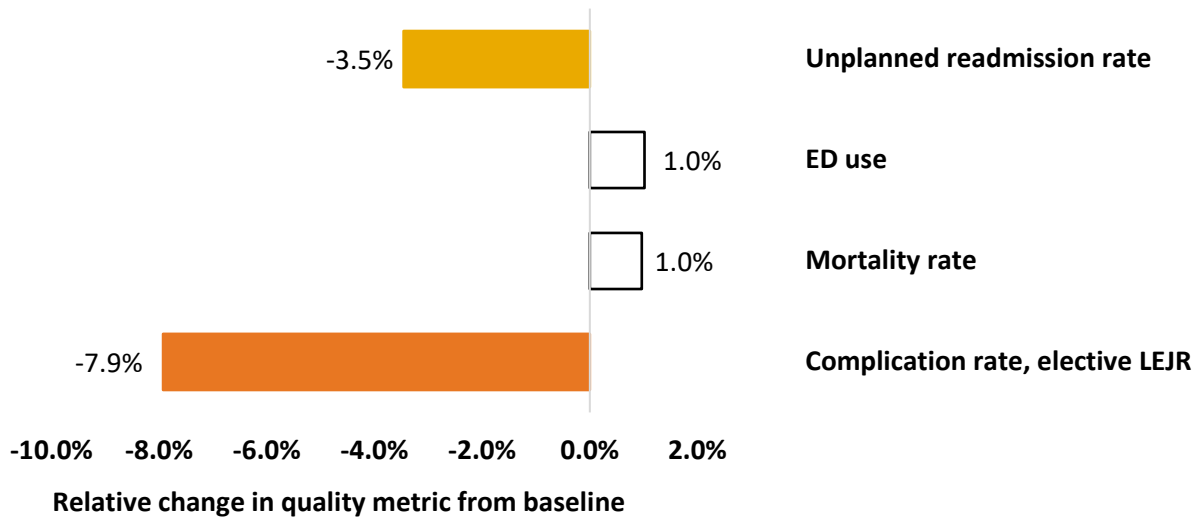
-3.5%   **Unplanned readmission rate**

1.0%   **ED use**

1.0%   **Mortality rate**

-7.9%   **Complication rate, elective LEJR**

-10.0%   -8.0%   -6.0%   -4.0%   -2.0%   0.0%   2.0%

Relative change in quality metric from baseline

Source: CJR evaluation team analysis of Medicare claims and enrollment data for episodes initiated in 2012 through 2014 that ended between April 2012 and March 2015 (baseline) and episodes initiated during or after April 2016 that ended by December 2019 (intervention).

Notes: The estimates in this exhibit are the result of a DiD model. DiD estimates that are significant at the 1%, 5%, or 10% significance level are indicated by red, orange, or yellow shaded bars, respectively.

The relative change from baseline is calculated as the DiD estimate as a percent of the CJR baseline level.

Because CJR participant hospitals shifted a lower share of TKAs to the hospital outpatient setting, the control group includes outpatient TKA episodes that have been weighted to balance the episode volume in the CJR hospitals.

DiD = difference-in-differences, ED = emergency department, LEJR = lower extremity joint replacement, PY = performance year, TKA = total knee arthroplasty.

d. Conclusion

Quality of care improved or was maintained according to the claims-based measures we examined for mandatory CJR hospitals. During the first four performance years, the CJR model resulted in relative reductions in the 90-day unplanned readmission rate and the elective LEJR complication rate. There was no change in 90-day ED use or the mortality rate.

5. What was the impact of the CJR model on functional status, pain, and care experiences?

The CJR model lowered average LEJR episode payments among mandatory CJR hospitals, primarily through reductions in institutional PAC (see Section II.A.1 and Section II.A.3). In addition to the claims-based measures of care quality assessed in Section II.A.4, reductions in PAC use may affect patient outcomes that cannot be assessed with Medicare claims data. Patients' functional status, pain, and care experience throughout the episode are of particular importance for understanding the impact of the CJR model on quality of care.

We therefore surveyed patients after the end of their episodes to learn about their experiences and long-term recovery. We also examined patient assessments conducted while patients were receiving PAC to understand patient function and pain in the period just following LEJR surgery.

---

a. Key Findings

- After the end of their LEJR episodes, CJR and control survey respondents on average had similar self-reported changes in functional status and pain.

- For patients discharged from the hospital to an IRF, SNF, or HHA, assessment data indicated that CJR and control patients had similar improvement in functional status while in the PAC setting.

- CJR and control survey respondents reported similar satisfaction with overall recovery, care management, and care transition experiences. However, CJR respondents required more help from caregivers when they returned home compared to control respondents.

- For survey respondents with a hip fracture there are indications of worse outcomes:

  - CJR respondents reported less improvement in functional status, specifically difficulty rising from sitting, standing, and using the toilet, and less satisfaction with treatment instructions than control respondents.

  - CJR respondents required more help from caregivers at home than control respondents.

---

b. Methods

Patient survey

We surveyed patients after the end of their inpatient LEJR episode to determine if CJR patients differed from control patients on several patient-

Acronyms	
ADLs	activities of daily living
BPCI	Bundled Payment for Care Improvement
DiD	difference-in-differences
HH	home health
HHA	home health agency
IRF	inpatient rehabilitation facility
IRF-PAI	inpatient rehabilitation facility-patient assessment instrument
LEJR	lower extremity joint replacement
MDS	minimum data set
MSA	metropolitan statistical area
OASIS	outcome and assessment information set
PAC	post-acute care
pp	percentage point
SNF	skilled nursing facility
TKA	total knee arthroplasty

reported outcomes.[37] Measures included change in functional status and pain (recalled from before their surgery to the time of the survey), as well as satisfaction with overall recovery, satisfaction with care management, experience with care transitions, and caregiver help needed after returning home. We estimated risk-adjusted differences between CJR and control respondents, accounting for beneficiary, hospital, and MSA attributes. See Appendix C and Appendix G for more detail on these methods.

The survey contained eight questions about dimensions of function and pain, including difficulty with rising, standing, walking, using stairs, and toileting; using a mobility aid (e.g., cane, walker); pain interfering with daily activities; and use of pain medication (Appendix H). For each of the eight questions, respondents were asked to rate function or pain using a Likert scale, which we transformed into percentage terms, where the numerator is the estimated difference in change (from before their surgery to after the end of the episode), and the denominator is the average recalled functional status among CJR respondents from before their surgery. Measures of satisfaction and caregiver help were scaled from 0 (worst outcome) to 100 (best outcome) points. Measures of experience with care transitions are reported in percentage terms (with differences interpreted on a pp scale). All outcomes in all domains are scaled such that higher values indicate more favorable outcomes. Therefore, positive differences indicate more favorable results for CJR respondents relative to control respondents, while negative differences indicate less favorable results for CJR respondents relative to control respondents.

Data were collected in two waves that covered episodes with inpatient discharges from mandatory CJR hospitals in March or April 2019 and in September or October 2019. We excluded prior survey waves from this analysis due to substantive changes to the underlying composition of our sample between 2018 and 2019, as described in Appendix C. Our starting sample included all CJR episodes initiated during these periods (N=17,190) and all control episodes (N=17,314) (Exhibit 20).

Exhibit 20: Sample size and response rate overall and for patients with hip fracture, waves 5 and 6 combined

Group	Patients surveyed (starting sample)		Survey responses received (analytic sample)		Response rate		
	CJR	Control	CJR	Control	CJR	Control	p-value
Overall[a]	17,190	17,314	11,273	11,765	65.6%	68.0%	p<0.05
Hip fracture	2,617	2,472	1,080	1,088	41.3%	44.0%	0.14

Source: CJR evaluation team analysis of survey data for patients with discharge from LEJR surgery in March, April, September, or October 2019.

Notes: Differences in CJR and control response rates that are significant at the 1%, 5%, or 10% significance level are indicated by red, orange, or yellow shaded cells, respectively.
LEJR = lower extremity joint replacement.
[a] Overall includes patients with elective inpatient LEJR or inpatient LEJR due to hip fracture.

---

[37]   The median time at which surveys were returned was 31 days after the conclusion of the patient's 90-day post-discharge period.

The overall response rate was 65.6% among CJR patients and 68.0% among control patients (significantly different at p<0.05). Survey results are based on 11,273 completed survey responses from CJR patients and 11,765 from control patients. Despite the differential nonresponse rates, CJR and control respondents had similar characteristics on average (Appendix I), suggesting that differences in non-response did not introduce differences in unobservable factors that would bias our results.

We separately analyzed responses of beneficiaries who received LEJR surgery after a hip fracture. The overall response rate among hip fracture patients was 41.3% among CJR patients and 44.0% among control patients (not statistically different between the two groups), yielding 1,080 and 1,088 total responses, respectively. CJR and control respondents with hip fractures generally had similar characteristics (Appendix I).

Assessment-based measures

We conducted a DiD analysis to estimate the differential change in functional status and pain measures for patients discharged from mandatory CJR hospitals to an IRF, SNF, or HHA. These measures derive from comprehensive assessments completed for every patient at the start and end of their PAC stay. Different assessment instruments are used in each PAC setting, as follows:

- Inpatient Rehabilitation Facility Patient Assessment Instrument (IRF-PAI) for IRF patients;

- Minimum Data Set (MDS) for SNF patients; and

- Outcome and Assessment Information Set (OASIS) for HHA patients.

The measures differ across these setting-specific instruments, as does the timing between the admission and discharge assessments.[38] While this precludes direct comparisons of patients across settings, the measures provide valuable information about changes in functional status and pain while a patient is receiving care in a particular setting. We focus on the first PAC setting to which patients were discharged (IRF, SNF, or HHA) and measure changes in patient functional status and pain. We report one functional status measure for patients initially discharged to an IRF, two functional status measures and one pain measure for those initially discharged to a SNF, and two functional status measures and one pain measure for those initially discharged to an HHA.

The results are risk-adjusted to control for functional status at the initiation of PAC, health care service use before the anchor hospitalization, and beneficiary, market, and hospital

---

[38] The IRF-PAI, MDS, and OASIS assessments use different measures to assess functional status. Although CMS added cross-setting measures of functional status to each of the three assessment instruments in October 2018, we did not use the new measures in this analysis because these measures were not available in the baseline. Estimating DiD analyses for these measures is therefore infeasible.

characteristics.[39] Risk adjustment is important because the CJR model has affected the initial discharge setting for LEJR patients, as well as the duration of PAC in each setting.

We conducted additional analyses to assess how the complexity of CJR patients discharged to an IRF, SNF, or HHA changed relative to control patients. We used a similar DiD design to estimate the unadjusted differential change in patient complexity measures obtained from IRF, SNF, and HHA admission assessments, claims, and enrollment data. Details about the DiD estimator and risk adjustment models are discussed in Appendix C (Section III.C).

The assessment-based analyses rely on the same baseline period as the claims-based analyses. The intervention period for the assessment analyses is one quarter shorter (from April 2016 to September 2019) because of the longer time needed for PAC assessment data to become available. The assessment-based analyses include only LEJRs performed in the inpatient setting because Medicare does not cover SNF care following outpatient procedures and outpatient TKA patients being first discharged to an IRF are rare. Results for mandatory CJR hospitals are reported in this section. We also report results for voluntary CJR opt-in and non-opt-in hospitals in Section II.A.7.

c. Results

All LEJR episodes

This section presents the patient survey results for all respondents and assessment results for patients at mandatory CJR hospitals discharged to either an IRF, SNF, or HHA. The analyses include both patients who received elective LEJR surgery and those who had surgery as a result of a fracture. Both the survey and assessment results measure functional status and pain. The survey results also examine CJR patients' satisfaction with their overall recovery and care management, their experience with care transitions, and amount of caregiver help that was needed after returning home.

Functional status and pain

Patient survey findings

Overall, respondents from mandatory CJR hospitals and control respondents reported similar improvements from before their surgery to after the episode on all eight measures of functional status and pain (Exhibit 21). Differences between CJR and control respondents in the amount of improvement are reported as means and as a percentage of the average status CJR respondents recalled prior to their hospitalization. These differences varied in direction and were not statistically significant, with the exception of one measure. The only statistically significant result was that CJR respondents reported less dependence on pain medication for the joint they had replaced (1.1%, $p<0.10$).

---

[39]   The pain measure for those initially discharged to a SNF was not risk adjusted, following the specifications of the MDS 3.0 Quality Measure for short-stay patients used in the CMS Nursing Home Five-Star Rating System.

Exhibit 21: CJR and control survey respondents experienced similar improvement in self-reported functional status and pain

Survey measure	Response range	Mean change in self-reported measure from before the hospitalization to after the episode *Higher value represents a more favorable change*		Difference between CJR and control groups (% difference)[a]
		CJR	Control group	
Ability to walk by yourself without resting	-4 to 4	0.76	0.74	0.02 (0.9%)
Difficulty walking up or down 12 stairs	-3 to 3	0.75	0.75	-0.01 (-0.2%)
Difficulty rising from sitting	-4 to 4	1.22	1.21	0.00 (0.1%)
Difficulty standing	-4 to 4	1.17	1.18	-0.01 (-0.4%)
Use of a mobility aid	-2 to 2	0.16	0.16	0.00 (0.1%)
Difficulty getting on/off the toilet	-4 to 4	1.36	1.37	-0.01 (-0.2%)
Frequency that pain interferes with normal activities	-4 to 4	1.98	1.97	0.01 (0.4%)
Medication use for pain in the joint you had replaced	-3 to 3	0.60	0.57	0.03 (1.1%)

Source: CJR evaluation team analysis of patient survey data for episodes with discharge in March, April, September, or October 2019.

Notes: The estimates in this exhibit are the results of a cross-sectional regression model, weighted for sampling and nonresponse. Estimates that are significant at the 1%, 5%, or 10% significance level are indicated by red, orange, or yellow shaded cells, respectively.

[a] The change in a given measure refers to the difference between a respondent's self-reported status at the time of the survey and the respondent's recalled status prior to the hospitalization. Estimated changes, and the difference between changes in the CJR and control groups, are reported in "level" terms (that is, levels of the Likert scale for each measure). Percentage differences are equal to the difference between CJR and control groups divided by the average CJR recalled status prior to the hospitalization.

Assessment-based results

Clinical assessments indicate that patients treated at mandatory CJR hospitals and control patients had similar improvements in functional status while receiving care in their first PAC setting after hospital discharge (Exhibit 22). The results for each PAC setting are presented in detail below.

Exhibit 22: The CJR model did not generally have an impact on improvements in functional status

First PAC setting	Measure	DiD	p-value
IRF	IRF mobility score[a]	-0.2	p=0.41
SNF	Improved mobility	0.6	p=0.64
	Improved toilet use	-3.3	p=0.12
	Without moderate to severe pain[b,c]	4.0	p<0.01
HHA	Improved ambulation/locomotion	0.2	p=0.75
	Improved bed transferring	-0.4	p=0.59
	Reduced pain	0.5	p=0.70

Source: CJR evaluation team analysis of Medicare claims and enrollment data, IRF-PAI data, MDS data, and OASIS data for episodes initiated in 2012 through 2014 that ended between April 2012 and March 2015 (baseline) and episodes initiated during or after April 2016 that ended by September 2019 (intervention).

Notes: The estimates in this exhibit are the result of a DiD model. DiD estimates that are significant at the 1%, 5%, or 10% significance level are indicated by red, orange, or yellow shaded cells, respectively.

The relative change from CJR baseline is calculated as the DiD estimate as a percent of the CJR baseline level.

ADLs = activities of daily living, DiD = difference-in-differences, HHA = home health agency, IRF = inpatient rehabilitation facility, IRF-PAI = Inpatient Rehabilitation Facility Patient Assessment Instrument, MDS = Minimum Data Set, OASIS = Outcome and Assessment Information Set, PAC = post-acute care, SNF = skilled nursing facility.

[a] The mobility score is a composite measure of related ADLs: ability to transfer from a bed to a chair, wheelchair, or standing; transfer on and off the toilet; walk or use a wheelchair; and navigate stairs. The mobility score ranges from 4 (total assistance) to 28 (complete independence). A positive change in mobility score from IRF admission to discharge indicates that a patient's mobility improved during their IRF stay.

[b] We cannot be certain that there is no impact of the model on the SNF pain measure because this outcome failed parallel trends tests (Appendix K). Parallel trends are an assumption that underlies our methodological approach, and without it we do not necessarily believe the control group acts as an accurate representation of what would have occurred in CJR hospitals in the absence of the CJR model. Please see Appendix C (Section III.C.1.c) for additional details.

[c] The pain measure for those initially discharged to a SNF was not risk adjusted following the specifications of the MDS 3.0 Quality Measure for short-stay patients used in the CMS Nursing Home Five-Star Rating System.

Patients first discharged to an IRF

Among patients first discharged to an IRF, the average complexity increased between the baseline and intervention periods for both CJR and control patients, but the increase was greater for CJR patients (Appendix J, Exhibit J-1).[40] Despite this relative increase in complexity, nearly all patients improved in mobility during their IRF stays, and the improvement was similar for CJR and control patients. The improvement in mobility during the

---

[40] As discussed in Section II.A.6b, throughout this annual report, we define patient complexity in terms of patient characteristics that are associated with higher episode payments.

IRF stay was greater in the intervention period than in the baseline period for both groups, with no differences between CJR and control patients (Appendix D, Exhibit D-2).[41]

Patients first discharged to a SNF

There was little evidence of a difference in the complexity of CJR patients first discharged to a SNF, relative to control patients (Appendix J, Exhibit J-2).

The proportion of CJR patients without moderate to severe pain at SNF admission increased by 17.7pp from baseline to intervention (from 58.3% to 76.0%) (Appendix D, Exhibit D-2). The proportion of control patients without moderate to severe pain increased by 13.8pp (from 53.2% to 67.0%). Thus, the proportion of CJR patients without moderate to severe pain increased by 4.0pp relative to control patients, or 6.8% from the CJR baseline (p<0.01). While this estimate is statistically significant, we cannot be certain that it is an impact of the CJR model because this outcome failed parallel trends tests (Appendix K, Exhibit K-2). Parallel trends is an assumption that underlies our methodological approach. Because failure to pass the parallel trends test indicates that CJR and control group hospitals may have had different patterns of moderate to severe pain at SNF admission in the baseline, we cannot isolate the impact of the CJR model on this outcome. Please see Appendix C (Section III.C.1.c) for additional details.

The CJR model had no effect on the proportion of patients who improved in mobility or in toileting during their SNF stays (Appendix D, Exhibit D-2).

Patients first discharged to an HHA

Among patients first discharged to an HHA, the average complexity increased between the baseline and intervention periods for both CJR and control patients (Appendix J, Exhibit J-3). In the intervention period, at least 80% of CJR and control patients improved during HH care in measures of ambulation/locomotion, bed transferring, and pain with activity. This improvement during HH care was greater than in the baseline period, with no differences between CJR and control patients (Appendix D, Exhibit D-2).

---

[41]   The mobility score is a composite measure of related activities of daily living (ADLs): ability to transfer from a bed to a chair, wheelchair, or standing; transfer on and off the toilet; walk or use a wheelchair; and navigate stairs. The mobility score ranges from 4 (total assistance) to 28 (complete independence). A positive change in mobility score from IRF admission to discharge indicates that a patient's mobility improved during their IRF stay.

Satisfaction with overall recovery and care management, experience with care transitions, and caregiver help

Patient survey findings

The patient survey also asked about satisfaction with overall recovery, satisfaction with care management, experience with care transitions, and caregiver help needed after returning home.[42] Patients self-reported each of these measures roughly three to four months after their LEJR.

Satisfaction with overall recovery and care management

CJR and control respondents were generally satisfied with overall recovery and care management, with average scores in these measures ranging from 78 out of 100 to 84 out of 100 (higher scores indicate greater satisfaction). There were no statistically significant differences between CJR and control respondents on any measure of satisfaction related to overall recovery or care management (Appendix I, Exhibit I-1).

Experience with care transitions

Both the CJR and control respondents generally indicated positive experiences with care transitions. Most respondents in both groups reported that they were discharged at the right time (88%), received the right amount of PAC (85%), and had access to all necessary durable medical equipment (over 90%). There were no statistically significant differences between the CJR and control respondents on any of these measures (Appendix I, Exhibit I-1).

Caregiver help

Approximately 95% of both CJR and control respondents had a caregiver from whom they could receive help after returning home (Appendix I, Exhibit I-1). Among those who received caregiver help at home, CJR respondents required more caregiver help than control respondents: a difference of 1.3 points based on a composite measure of help with three activities of daily living (ADLs), scaled from 0 to 100 (p<0.05, Exhibit 23). CJR respondents required more help from caregivers than did control respondents with each of the three ADLs: putting on and taking off clothes (-1.7 points, p<0.10), bathing (-1.1 points, p=0.101), and using the toilet (-1.1 points, p<0.05). These differences could be interpreted as roughly one to two additional CJR respondents out of 100 needing more caregiver help with one or more of these ADLs than control respondents.

---

[42]   Measures of satisfaction and caregiver help are scaled from 0 to 100 points. Measures of experience with care transitions are binary measures that are presented as percent of respondents in agreement (ranging from 0 to 100 percent with differences expressed in pp).

Exhibit 23:   CJR survey respondents required more help from caregivers at home than control respondents

Source: CJR evaluation team analysis of patient survey data for episodes with discharge in March, April, September, or October 2019.

Notes: Differences that are significant at the 1%, 5%, or 10% significance level are indicated by red, orange, or yellow shading, respectively.

Measures of caregiver help required among respondents who received any help are scaled from 0 to 100 points, where 0 = "complete help needed," 50 = "some help needed," and 100 = "no help needed."

The composite summarizes the amount of help needed across all three activities of daily living. Differences between CJR and control responses are reported in point terms.

Hip Fractures

Patient survey results

In addition to the main survey analysis, which included both elective inpatient LEJR and inpatient LEJR due to hip fracture at mandatory CJR hospitals (and control hospitals), we conducted a separate analysis focused on survey respondents with hip fractures. Surgeries for hip fractures are not elective (not planned ahead of time) and patients' conditions are generally more severe than for those having elective surgeries, in terms of decline in physical function and length of recovery.[43,44] Therefore, beneficiaries with hip fractures may be more sensitive to changes in care that occur under the CJR model (e.g., decreased use of institutional PAC or

---

43   Le Manach Y, Collins G, Bhandari M, et al. Outcomes After Hip Fracture Surgery Compared With Elective Total Hip Replacement. *JAMA*. 2015;314(11):1159–1166. doi:10.1001/jama.2015.10842

44   Wu VJ, Ross BJ, Sanchez FL, Billings CR, Sherman WF. Complications Following Total Hip Arthroplasty: A Nationwide Database Study Comparing Elective vs Hip Fracture Cases. J Arthroplasty. 2020 Aug;35(8):2144-2148.e3. doi: 10.1016/j.arth.2020.03.006. Epub 2020 Mar 10. PMID: 32229152.

improved care coordination). We did not conduct separate analyses of PAC assessment data for patients with hip fractures.

Functional status and pain

CJR respondents with hip fractures experienced less improvement than control respondents with hip fractures, from before their surgery to after the end of the episode, on three of eight measures of functional status and pain (Exhibit 24). CJR respondents reported less improvement than control respondents in rising from sitting (-2.2%, $p<0.05$), standing (-3.6%, $p<0.01$), and using the toilet (-5.0%, $p<0.01$). For each of these three measures, more than 60% of respondents in both the CJR and control groups regained or exceeded their pre-hospital function. However, for each of the three measures, the difference equates to roughly three to six more CJR respondents out of 100 with hip fractures indicating decline in functional status from before their surgery (i.e., before their fracture) to the end of their episode (Appendix I, Exhibit I-5). Differences between CJR and control respondents were not statistically significant for the other five measures, although the CJR respondents had point estimates indicating less improvement than control respondents in four of the five measures (ranging from -0.4% to -3.0%).

Exhibit 24: CJR survey respondents with hip fractures reported less improvement in functional status than control respondents

Survey measure	Response range	Mean change in self-reported measure from before the hospitalization to after the episode[a] *Higher value represents a more favorable change*		Difference between CJR and control respondents (% difference)[b]
		CJR	**Control**	
Ability to walk by yourself without resting	-4 to 4	-0.72	-0.71	-0.01 (-0.4%)
Difficulty walking up or down 12 stairs	-3 to 3	-0.58	-0.50	-0.08 (-3.0%)
Difficulty rising from sitting	-4 to 4	-0.37	-0.29	-0.08 (-2.2%)
Difficulty standing	-4 to 4	-0.34	-0.20	-0.14 (-3.6%)
Use of a mobility aid	-2 to 2	-0.61	-0.60	-0.01 (-0.6%)
Difficulty getting on/off the toilet	-4 to 4	-0.17	0.03	-0.20 (-5.0%)
Frequency that pain interferes with normal activities	-4 to 4	-0.39	-0.37	-0.02 (-0.4%)
Medication use for pain in the joint you had replaced	-3 to 3	-0.31	-0.32	0.01 (0.1%)

Source: CJR evaluation team analysis of patient survey data for patients with a hip fracture discharged in March, April, September, or October 2019.

Notes: The estimates in this exhibit are the result of a cross-sectional regression model, weighted for sampling and nonresponse. Estimates that are significant at the 1%, 5%, or 10% significance level are indicated by red, orange, or yellow shaded cells, respectively.

[a]   Pre-hospital functional status refers to the respondents' function prior to the fracture. This explains why, on average, respondents report worse outcomes at the time of the survey than prior to their hospitalization.

[b]   The change in a given measure refers to the difference between a respondent's self-reported status at the time of the survey and the respondent's recalled status prior to the hospitalization. Estimated changes, and the difference between

changes among CJR and control respondents, are reported in "level" terms (that is, levels of the Likert scale for each measure). Percentage differences are equal to the difference between CJR and control respondents divided by the average CJR recalled status prior to the hospitalization.

Satisfaction with overall recovery and care management, experience with care transitions, and caregiver help[45]

Satisfaction with overall recovery and care management

Among CJR and control respondents with hip fractures, there was no difference in satisfaction with overall recovery (see Appendix I, Exhibit I-2). CJR respondents with a hip fracture reported less satisfaction with treatment instructions (-2.5 points, p<0.10; Exhibit 25) than did control respondents, but there were no other statistically significant differences between CJR and control respondents in measures of satisfaction with care management.

Exhibit 25:   CJR survey respondents with hip fractures were less satisfied with treatment instructions than control respondents with hip fractures
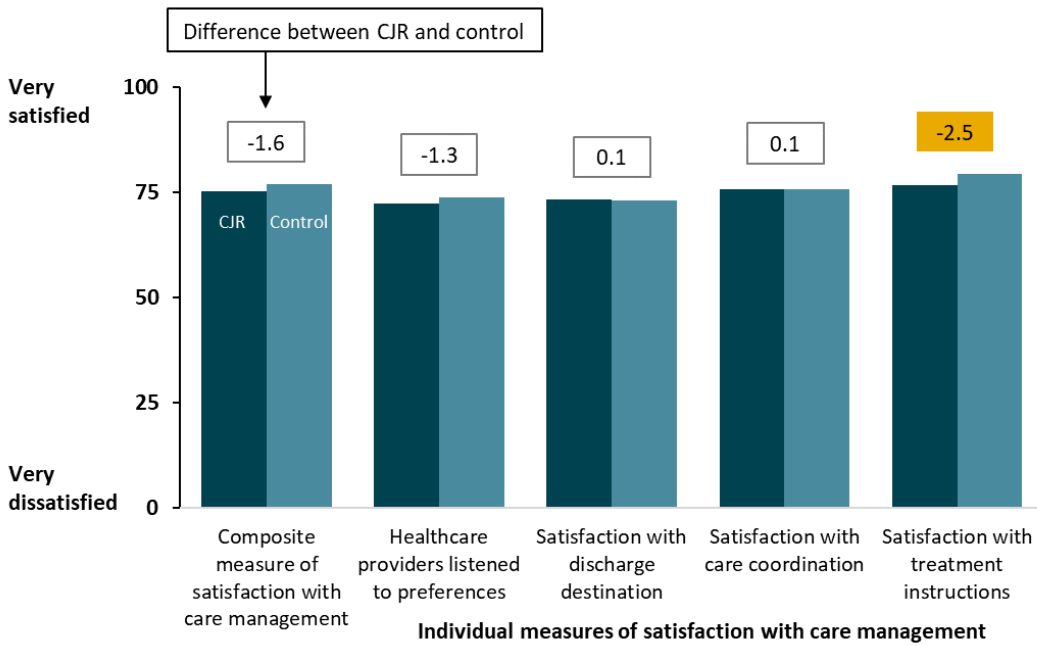
Source: CJR evaluation team analysis of patient survey data for episodes with discharge in March, April, September, or October 2019.

Notes: Differences that are significant at the 1%, 5%, or 10% significance level are indicated by red, orange, or yellow shading, respectively.

Measures of satisfaction with care management are scaled from 0 to 100 points, where 0 = "very dissatisfied," 25 = "somewhat satisfied," 50 = "neither satisfied nor dissatisfied," 75 = "somewhat satisfied," and 100 = "very satisfied"

The composite summarizes the overall satisfaction across all four measures of care management. Differences between CJR and control responses are reported in point terms.

---

[45]   Measures of satisfaction and caregiver help are scaled from 0 to 100 points. Measures of experience with care transitions are binary measures ranging from 0 to 100 percent (with differences expressed in pp).

Experience with care transitions

Among respondents with hip fractures, 77 out of 100 indicated that they received the right amount of post-acute care, versus 80 out of 100 control respondents, a significant difference of 3.0pp (p<0.05; Exhibit 26). There were no differences in the proportion of respondents who reported that they were discharged from the hospital at the right time, or that they had access to all necessary durable medical equipment at home.

Exhibit 26: Among respondents with hip fractures, care transition experiences were generally positive, but slightly fewer CJR respondents reported that they received sufficient post-acute care
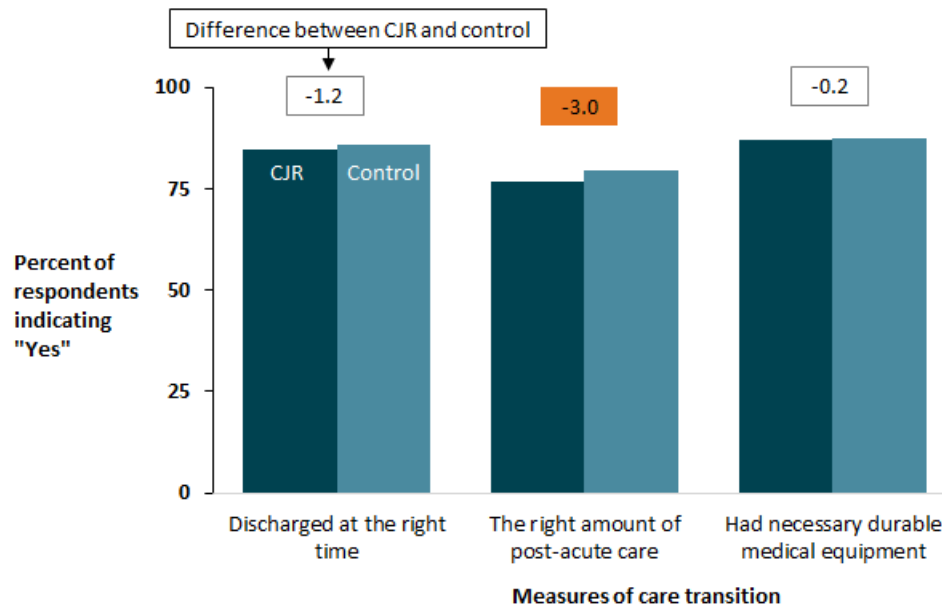
Source: CJR evaluation team analysis of patient survey data for episodes with discharge in March, April, September, or October 2019.

Notes: Differences that are significant at the 1%, 5%, or 10% significance level are indicated by red, orange, or yellow shading, respectively.

All outcomes are in percentages, ranging from 0 to 100. Differences between CJR and control responses are reported in pp.

Caregiver help

Approximately 96% of CJR and control survey respondents with a hip fracture had a caregiver from whom they could receive help after returning home (see Appendix I, Exhibit I-2). Among those who received caregiver help at home, CJR respondents required more caregiver help than control respondents: a difference of 5.6 points based on a composite of three ADLs scored from 0 to 100 (p<0.01; Exhibit 27). CJR respondents reported needing more help than control respondents in putting on or taking off clothes (-5.3 points, p<0.05), bathing (-6.2 points, p<0.01), and using the toilet (-6.2 points, p<0.01). These differences can be interpreted as roughly five to 10 additional CJR respondents out of 100 requiring more caregiver help with at least one of these ADLs relative to control respondents.

Exhibit 27: **CJR survey respondents with hip fractures required more help from caregivers at home than control respondents with hip fractures**
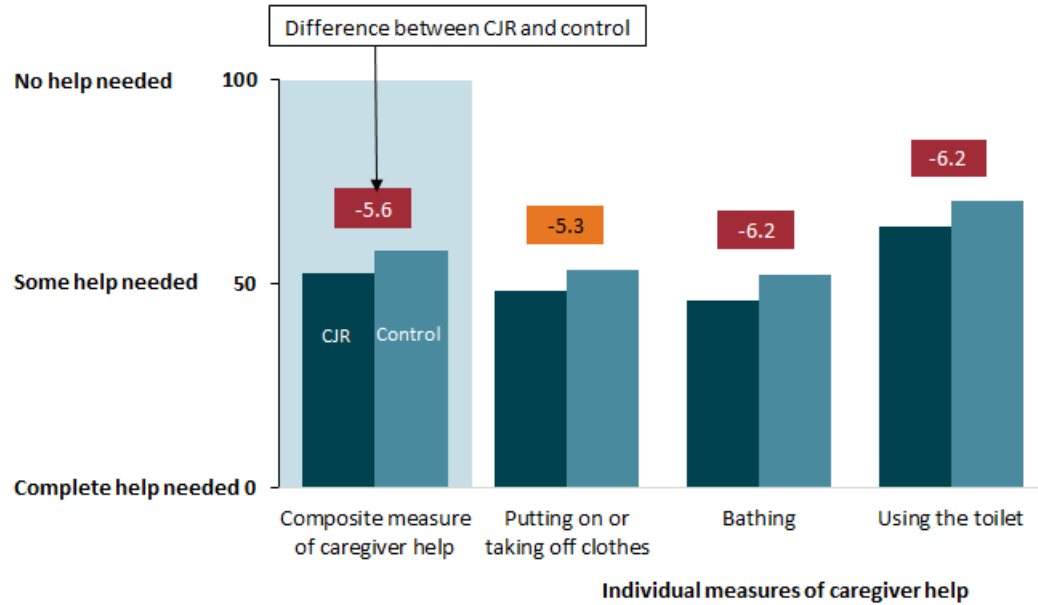
Source: CJR evaluation team analysis of patient survey data for episodes with discharge in March, April, September, or October 2019.

Notes: Differences that are significant at the 1%, 5%, or 10% significance level are indicated by red, orange, or yellow shading, respectively.

Measures of caregiver help required among respondents who received any help are scaled from 0 to 100 points, where 0 = "complete help needed," 50 = "some help needed," and 100 = "no help needed."

The composite summarizes the amount of help needed across all three activities of daily living. Differences between CJR and control responses are reported in point terms.

Sensitivity findings

We conducted additional analyses to understand the extent to which overlap from BPCI Advanced may have influenced our results. As discussed in Section II.A.1.b, hospital participation in BPCI Advanced differs between the CJR and control groups because the CJR model takes precedence over the BPCI Advanced model and CJR hospitals are not eligible to participate in BPCI Advanced for LEJR. Overall, roughly 46% of control respondents across all LEJR episodes were treated by BPCI Advanced participants (36% of control respondents with hip fractures were treated by BPCI Advanced participants). This may bias our results if outcomes from hospitals in BPCI Advanced are different than they would have been in the absence of BPCI Advanced. To explore this concern, we replicated each of the survey analyses, for all inpatient LEJRs and for surgeries occurring due to hip fracture, excluding any control respondents treated by a BPCI Advanced participant. We report these results in Appendix I, Exhibits I-3 and I-4. Results were substantively the same, indicating that potential overlap from BPCI Advanced is not materially influencing our results. In particular, overlap with BPCI Advanced does not account for differences in functional status among CJR and control respondents with hip fractures.

d. Conclusion

When considering all survey respondents (elective LEJR and LEJR due to hip fracture) discharged from mandatory CJR participant hospitals, the CJR model had no adverse effect on patient function or pain. Among all survey respondents, CJR and control respondents had similar improvement in functional status and pain from the week before their surgery to after the end of the episode when they completed the survey. PAC assessment analyses indicate that CJR and control patients had similar improvements in functional status and pain while receiving care in their first PAC setting.

CJR and control survey respondents reported similar satisfaction with overall recovery, care management, and care transitions. When returning home after surgery, CJR respondents required more help from caregivers with activities of daily living than did control respondents. While this result is consistent with the reduction in PAC use under the CJR model, our analysis is not designed to definitively conclude whether changes in PAC use were the reason that CJR respondents required more assistance at home.

The lack of the CJR model's impact on functional status and pain measures in both survey and assessment data is consistent with results for LEJRs at mandatory CJR hospitals in the third annual evaluation report, which covered LEJRs in 2018 (including two waves of survey data collection spanning March, April, August, and September 2018). Likewise, the preceding evaluation report indicated similar satisfaction between CJR and control survey respondents, and greater reliance on caregiver help among CJR respondents relative to control respondents.

Among survey respondents with hip fractures, CJR respondents reported less improvement in functional status than control respondents, despite robust controls for multiple measures of patient complexity and hospital- and market-level factors (see Appendix C, Section VI.D for more detail). Specifically, CJR respondents reported less improvement in rising from sitting, standing, and getting on or off the toilet. For each of these three measures, more than 60% of respondents in both the CJR and control groups regained or exceeded their pre-hospital function. However, for each measure, the differences equate to roughly three to six additional CJR respondents with hip fractures out of every 100 who declined in function from before their fracture until after the end of the episode. CJR survey respondents with hip fractures also required more help from caregivers at home than control respondents with hip fractures. These results remained after excluding all BPCI Advanced LEJR episodes from the control group (roughly one-third of control respondents), signifying that differences were not attributable to any overlap from BPCI Advanced. These findings are also similar to previous results from the two survey waves conducted in 2018 (Appendix I, Exhibit I-6), which indicated that CJR respondents with hip fractures had worse changes in function for six of eight measures, although only one difference (difficulty standing) was significant at the 10% level. Despite differences in functional status and the need for caregiver help, there was no difference between CJR and control respondents with hip fractures in satisfaction with overall recovery. Patients with hip fractures may be particularly vulnerable to changes in health care services to reduce payments,

including changes to PAC use. We will continue to survey patients about their LEJR recovery and evaluate the impact of the CJR model on patients with fracture.

These results suggest that reductions in institutional PAC under the CJR model did not adversely affect functional status or pain among LEJR patients overall, which is consistent with prior evaluation results. Across all survey respondents, the CJR and control respondents reported similar changes in functional status and pain, and the PAC assessement data analysis showed no adverse impacts of CJR on functional status. However, among the subset of survey respondents with a hip fracture, CJR respondents reported less improvement in functional status than control respondents.

6. *Did the model result in any unintended consequences?*

a. *Key Findings*

- The CJR model had no statistically significant impact on the total volume of elective LEJR discharges.

- For the highest volume and least complex episode group, elective Medicare Severity-Diagnosis Related Group (MS-DRG) 470, the CJR patient population was relatively healthier in the intervention period than in the baseline period. This relative change in patient complexity could make it easier for mandatory CJR hospitals to achieve payment and quality targets and thus receive reconciliation payments.

- For the other episode groups – elective MS-DRG 469, fracture MS-DRG 470, and fracture MS-DRG 469 – the CJR patient population did not change relative to the patient population at control hospitals.

- The CJR model likely had no impact on payments for services provided in the 30 days following the end of the episode.

6a. *What was the impact of the CJR model on total market volume of elective LEJR discharges?*

The mandatory CJR model targets LEJR surgery in part because of its prevalence in the Medicare population, with more than 400,000 surgeries on Medicare beneficiaries in 2014 and growth projected to continue.[46] LEJR volume has been trending upward since the 1990s, with rates of total hip or knee replacements approximately doubling among those 45 and older

---

[46]   Centers for Medicare & Medicaid Services. *Comprehensive Care for Joint Replacement Payment Model for Acute Care Hospitals Furnishing Lower Extremity Joint Replacement Services*; Final Rule 2015:1–282.

between 2000 and 2010.[47,48] Exhibit 28 shows the elective LEJR surgery rate for the Medicare FFS population from 2007.

Exhibit 28:   Increasing national trend in the elective LEJR discharge rate since 2007
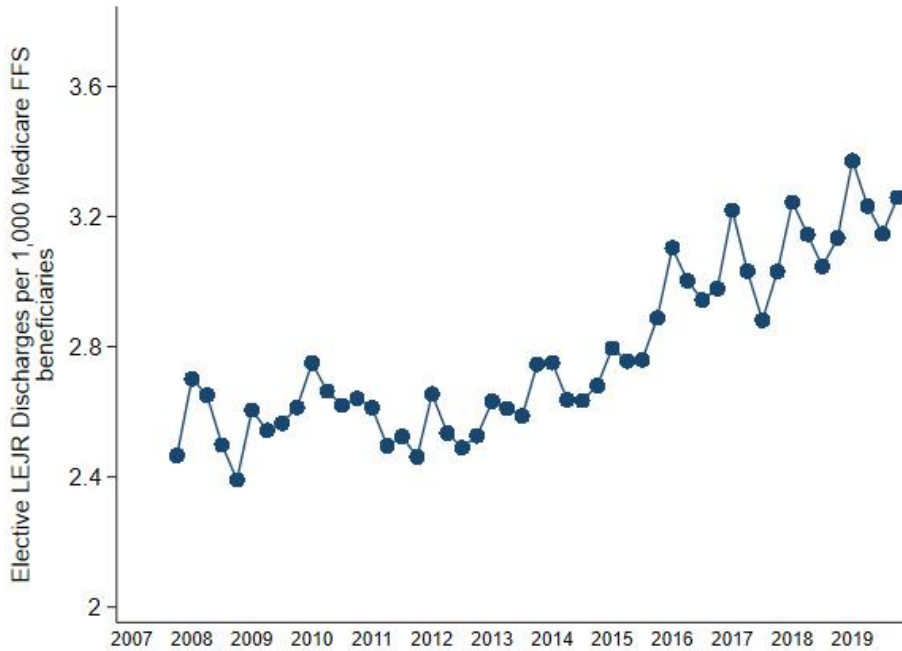
Source: CJR evaluation team analysis of Medicare claims and enrollment data for surgeries from October 2007 through December 2019.

Notes: Surgeries from 2018 and 2019 include LEJRs performed in inpatient and outpatient hospital settings.
FFS = fee-for-service, LEJR = lower extremity joint replacement.

In the context of overall growth in LEJRs, there are concerns that the CJR model itself could boost LEJR volume beyond what it would have been by making the surgery more financially rewarding to participant hospitals. Participant hospitals may be able to reduce average episode payments by providing elective LEJR to relatively healthier beneficiaries who otherwise would have foregone or delayed the procedure and are likely to have less costly episodes. Medicare savings due to the CJR model would be offset by the payments for these additional episodes.

CJR participant hospitals may also increase their volume of LEJR episodes if they shift surgeries from other hospitals through enhanced marketing, higher quality, or new gainsharing agreements

---

[47]   Wolford ML, Palso K, Bercovitz A. *Hospitalization for total hip replacement among inpatients aged 45 and over: United States*, 2000–2010. NCHS data brief, no 186. Hyattsville, MD: National Center for Health Statistics. 2015.

[48]   Williams SN, Wolford ML, Bercovitz A. *Hospitalization for total knee replacement among inpatients aged 45 and over: United States, 2000–2010*. NCHS data brief, no 210. Hyattsville, MD: National Center for Health Statistics. 2015.

with referring physicians. Shifts in volume across providers, however, would likely not have much effect on Medicare savings due to the CJR model.[49]

a. Methods

We analyzed the impact of the CJR model on market-level volume of elective LEJR surgeries. We examined volume in the market because sampling for participation in the model occurred at the MSA level and all acute care hospitals paid under the Medicare IPPS in the MSA, with few exceptions, were required to participate. This analysis compared all elective LEJRs in the 34 mandatory CJR MSAs relative to control group MSAs, weighted based on the probability that the

Acronyms	
BPCI	Bundled Payments for Care Improvement
CPT	current procedural terminology
FFS	fee-for-service
IPPS	inpatient prospective payment system
LEJR	lower extremity joint replacement
MSA	metropolitan statistical area
MS-DRG	Medicare Severity-Diagnosis Related Groups
PY	performance year
TKA	total knee arthroplasty

MSAs were selected into the mandatory group.[50] Elective LEJRs included: 1) LEJRs performed as hospital inpatient procedures (discharged under MS-DRG 469 or 470 and without fracture diagnosis codes); and 2) starting in 2018, TKA procedures performed in a hospital outpatient department (Current Procedural Terminology (CPT) code 27447). We included outpatient TKAs to ensure we were evaluating total volume of LEJRs in the market. See the third annual CJR evaluation report for additional discussion about outpatient TKA.[51] We examined LEJRs due to fracture separately because participants are unlikely to be able to affect volume of these surgeries.

We analyzed the impact of the CJR model on the volume of elective LEJRs in a market by estimating the relationship between CJR "dose" and the change in the elective LEJR rate (LEJRs per 1,000 Medicare FFS beneficiaries) in MSAs. Because the BPCI initiative provided similar incentives for boosting volume of elective LEJRs as the CJR model, we also included a BPCI dose for each MSA.

The CJR and BPCI doses were defined as the baseline market share of providers that ever participated in the CJR model or the risk-bearing phase of the BPCI initiative, respectively. The

---

[49] The effect on Medicare savings of shifts in volume across providers would depend on the difference in episode payments between the providers.

[50] All elective LEJRs for Medicare FFS beneficiaries in an MSA were included, whether they occurred at a CJR hospital or non-participating hospital.

[51] Centers for Medicare & Medicaid Services. Comprehensive care for joint replacement model - third annual report. https://innovation.cms.gov/data-and-reports/2020/cjr-thirdannrpt. 2020:116-121

baseline period for measuring dose was from October 2009 through September 2012, the three years before the first BPCI intervention period.[52]

We defined three CJR intervention periods:

- **Interim** begins the quarter that the CJR model was announced (July 2015) and ends the quarter before it was implemented (March 2016).[53]

- **PY1-2** begins the quarter that the CJR model took effect (April 1, 2016) and ends with the last quarter of PY2 (December 2017).

- **PY3-4** begins the quarter that changes to the number of mandatory CJR MSAs took effect (reduced from 67 to 34 mandatory MSAs) and outpatient TKA was removed from the inpatient only list (January 1, 2018) and ends with the last quarter of PY4 (December 2019).

We interacted each of the intervention period variables with the CJR dose and the BPCI dose. We also controlled for market-level characteristics, market and quarterly fixed effects, and a market-specific linear time trend. We report differences in surgeries rates between CJR and control group MSAs for the three different intervention periods and cumulatively across the four CJR performance years. We also tested whether the results were significantly different across intervention periods. See Appendix C (Section VII) for a full description of the methodology.

b. Results

There was no statistically significant difference in the volume of elective LEJRs between mandatory CJR MSAs and control group MSAs. In the first four performance years of the CJR model, the elective LEJR rate decreased by 0.08 per 1,000 Medicare FFS beneficiaries (p=0.21, Exhibit 29). This result was stable over time with similar differences in surgery rates between CJR and control group MSAs in the interim period (-0.04), PY1-2 (-0.06), and PY3-4 (-0.09). The differences in rates for each intervention period were not statistically significantly different from one another (p=0.65). As expected, we observed no change in the rate of LEJRs due to fracture.[54]

---

[52] Using the period prior to the intervention avoids circularity that would result from using LEJR market-quarter volume as both a component of the dependent variable and as a component of the exposure variable.

[53] The interim period used for the other impact analyses begins one quarter before the interim period used in the volume analysis to align the baseline period with the one set forth in the CJR model final rule.

[54] The estimated change in the fracture LEJR rate in the first four performance years due to the CJR model was a decrease of 0.01 per 1,000 FFS Medicare beneficiaries (p=0.52).

Exhibit 29: The CJR model did not have a statistically significant impact on the volume of elective LEJR surgeries in mandatory MSAs

Period	Predicted CJR MSA surgery rate (per 1,000 FFS beneficiaries)	Predicted control group MSA surgery rate (per 1,000 FFS beneficiaries)	Difference in surgery rates (per 1,000 FFS beneficiaries)	p-value
Interim	2.81	2.85	-0.04	0.33
PY1-2	2.90	2.96	-0.06	0.24
PY3-4	3.06	3.15	-0.09	0.24
PY1-4	2.98	3.06	-0.08	0.21

Source: CJR evaluation team analysis of Medicare claims and enrollment data for surgeries from October 2007 through December 2019.

Notes: R-squared = 0.94.

Estimates that are significant at the 99%, 95%, or 90% significance level are indicated by red, orange, or yellow shaded cells, respectively.

Interim includes surgeries from July 2015 to March 2016. PY1-2 includes surgeries from April 2016 through December 2017. PY3-4 includes surgeries from January 2018 through December 2019.

FFS = fee-for-service, LEJR = lower extremity joint replacement, MSA = metropolitan statistical area, PY = performance year.

c. Conclusion

There was no statistically significant difference in the volume of elective LEJRs between mandatory CJR MSAs and control group MSAs in the first four performance years. In future reports, we will continue to monitor any changes in elective volume as Medicare coverage of THA expands to include the hospital outpatient department and TKA is covered in ambulatory surgery centers beginning with PY5, as well as with the inclusion of outpatient LEJR as episodes in the CJR model starting in PY6. Additionally, we will adjust our analysis appropriately in PY5 to account for the impact of the COVID-19 pandemic on elective LEJR volume.

6b. Were there any indications that the CJR patient population was different in the intervention period than in the baseline period?

One potential unintended consequence of the CJR model would be if it influences hospital participants to select less complex patients. Less complex patients may require fewer resources and, therefore, have less costly episodes. In addition, treating less complex patients may make it easier for hospitals to achieve higher quality scores, which would lower the effective CMS discount and increase the quality-adjusted target price. As a result, participant hospitals could receive reconciliation payments because they treated a less complex mix of patients instead of lowering payments and improving quality through care redesign. Changes in a hospital's LEJR patient population could be due to factors outside of its control or due to intentional patient selection. Regardless of the reason for any change in patient complexity, the CJR model is intended to financially reward hospitals that lower episode payments or improve quality through care redesign, not changes in patient mix. Further, if the CJR model financially

rewards hospitals that have less complex patient populations, it could reduce access to care for more complex patients who require more resource intensive treatment.

The CJR model was designed to limit patient selection and appropriately account for patient complexity. First, all hospitals in selected MSAs are participating in the model, which limits the ability to selectively admit patients. Second, target prices differ by four episode groups determined by fracture status (elective or fracture) and MS-DRG (470 or 469) (Exhibit 30), which is intended to account for differences in patient resource needs. As a result, for changes in patient mix to affect a CJR hospital's reconciliation payments, the complexity of patients within an episode group would need to change.

Exhibit 30: There are quality-adjusted target prices for four episode groups, indicated by presence or absence of a fracture and MS-DRG
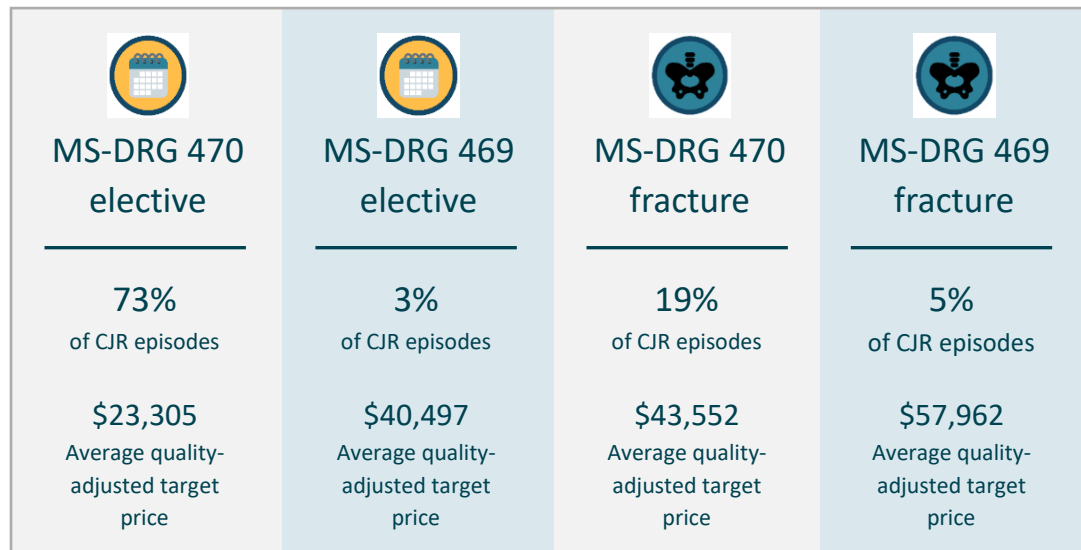
MS-DRG 470 elective	MS-DRG 469 elective	MS-DRG 470 fracture	MS-DRG 469 fracture
73% of CJR episodes	3% of CJR episodes	19% of CJR episodes	5% of CJR episodes
$23,305 Average quality-adjusted target price	$40,497 Average quality-adjusted target price	$43,552 Average quality-adjusted target price	$57,962 Average quality-adjusted target price

Source: CJR evaluation team analysis of Medicare claims and enrollment data for episodes initiated during or after April 2016 that ended by December 2019 (intervention) and analysis of CJR payment contractor target price data for PY1-4.

Notes: Reported shares and means are averages over the four performance years for mandatory CJR hospitals.
          MS-DRG = Medicare Severity-Diagnosis Related Group, PY = performance year.

There are several ways that patient complexity within an episode group could change. Hospitals could upcode the most complex patients in MS-DRG 470 (without major complications and comorbidities) to MS-DRG 469 (with major complications and comorbidities). If the hospital more completely documented conditions that could qualify as major complications or comorbidities to increase the number of patients coded as MS-DRG 469 instead of 470, then the complexity of patients in both groups would decrease, but the target prices would remain the same.[55] As a result, it would be easier for the hospital to have average episode payments in both groups that were below its target price. Another way that patient complexity could change is if

---

[55]   Recall a hospital's target prices are a changing blend of their own historical average payments and the historical regional average.

the number of healthier beneficiaries having LEJRs increased, which would reduce the average complexity of patients receiving LEJRs while increasing the volume of LEJR procedures. Similarly, a hospital's patient complexity could decrease if more complex Medicare beneficiaries were less likely to have LEJR procedures.

In this section, we evaluate if the patient population at mandatory CJR hospitals changed from the baseline to the intervention period. We compare CJR hospital patient characteristics associated with higher episode payments to the same patient characteristics at control hospitals to determine if any are unique to mandatory CJR hospitals.[56] This analysis allows us to evaluate relative changes in patient complexity that may be the result of the CJR model.

a. Methods

We analyzed changes in the mix of patients at mandatory CJR hospitals in two ways. Our first method focused on average episode payments, a composite measure of patient complexity, for each of the four episode groups. We isolated the impact of the CJR model on average episode payments that is associated with relative changes in the patient populations at mandatory CJR hospitals. This allowed us to estimate the specific impact of any patient mix changes on episode payments for each episode group, while controlling for hospital and market-level factors. (See Appendix C Section IX for details of this analysis.)

Acronyms	
ACH	acute care hospital
CI	confidence interval
DiD	difference-in-differences
ED	emergency department
ESRD	end-stage renal disease
HCC	hierarchical condition category
HH	home health
IRF	inpatient rehabilitation facility
LEJR	lower extremity joint replacement
LTCH	long-term care hospital
MSA	metropolitan statistical area
MS-DRG	Medicare Severity Diagnosis Related Group
pp	percentage point
PY	performance year
SNF	skilled nursing facility
TKA	total knee arthroplasty

In our second method, we evaluated changes in patient characteristics in each of the four episode groups to determine if the CJR patient population was different in the intervention period than in the baseline period. We examined changes in age, sex, race, Medicaid eligibility, disability status, health status, and prior health care use for inpatient LEJR patients from the baseline to the intervention period for CJR patients relative to control patients.

b. Results

Analyses of composite measure of patient complexity. As previously reported, LEJR patient complexity increased in both CJR and control hospitals.[57] Relative to the control group, however, the CJR patient population in the elective MS-DRG 470 episode group was less complex in the

---

[56] Associating complexity with higher episode payments is consistent with the adjustments to Medicare Advantage plan payments. Medicare Advantage payments are lower to plans with a relatively healthier mix of patients.

[57] Centers for Medicare & Medicaid Services. Comprehensive care for joint replacement model - third annual report. https://innovation.cms.gov/data-and-reports/2020/cjr-thirdannrpt. 2020:98-106

intervention period than the baseline. Because of CJR patient complexity changes, total payments decreased by $170 per episode for mandatory CJR hospitals relative to control hospitals (p<0.01, Exhibit 31). This $170 reduction in average total payments is due to changes in patient characteristics that we control for through risk adjustment, so it does not contribute to the estimated impact of the CJR model on episode payments for all LEJR episodes reported earlier (Section II.A.1). Our DiD impact estimates do not include the change in payments due to relative changes in patient complexity because they are intended to measure the impact of the CJR model on episode payments for similar patients. However, this analysis of unintended consequences indicates that episode payment changes differed between CJR and control hospitals because relative patient complexity decreased for CJR hospitals. As a result, some of the reconciliation payments made to CJR hospitals may be due to their decrease in patient complexity, which was not accounted for in the reconciliation payment calculation.

We examined whether coding changes that would shift patients from the less complex MS-DRG 470 to the more complex MS-DRG 469 could explain the relative reduction in patient complexity in the MS-DRG 470 episode group. For the elective MS-DRG 469 episode group, we found no evidence of a relatively less complex patient population at CJR hospitals (Exhibit 31). This, along with no change in the total market volume of elective LEJR discharges (Section II.A.6a), suggests that mandatory CJR hospitals are not shifting patients from MS-DRG 470 to MS-DRG 469.

Exhibit 31: Average episode payments for the elective MS-DRG 470 episode group decreased due to a less complex mix of patients under the CJR model
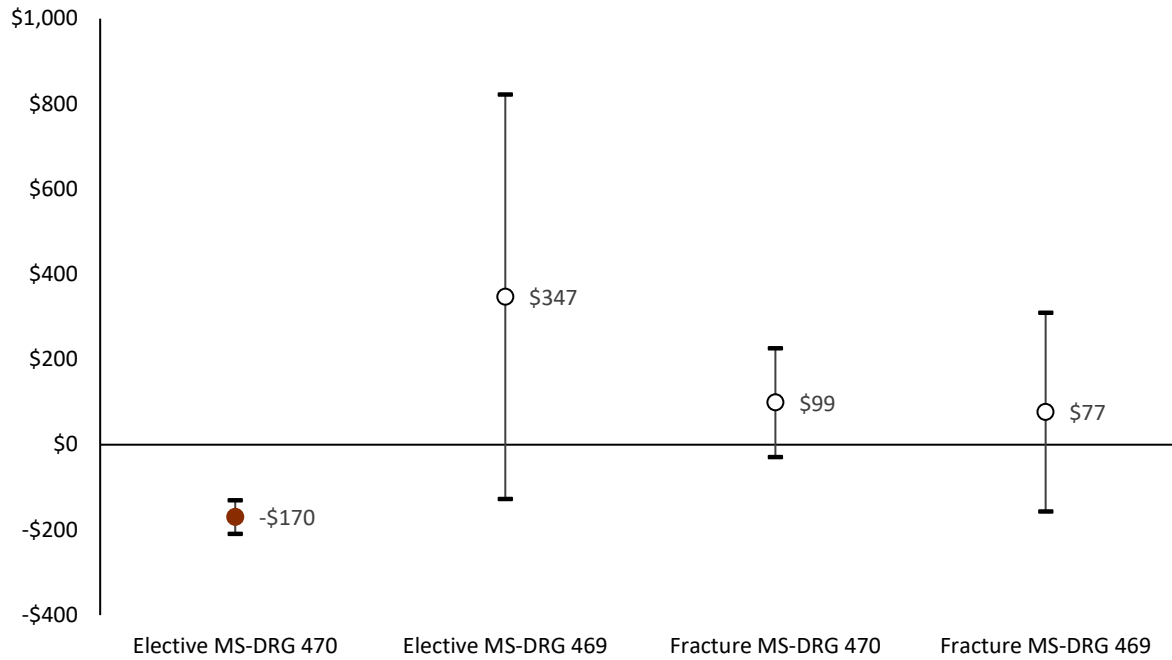
Source: CJR evaluation team analysis of Medicare claims and enrollment data for episodes initiated in 2012 through 2014 that ended between April 2012 and March 2015 (baseline) and episodes initiated during or after April 2016 that ended by December 2019 (intervention).

Notes: 90% CIs are plotted as vertical bars for relative changes in total episode payments that resulted from changes in patient mix in CJR hospitals relative to control hospitals. Each estimate is obtained from a separate analysis that measures how much of the relative change in total payments between CJR and control hospitals over the intervention period is attributable to relative changes in patient characteristics for the respective episode group.

Estimates that are significant at the 1%, 5%, or 10% significance level are indicated by red, orange, or yellow shaded points, respectively.

CI = confidence interval, MS-DRG = Medicare Severity-Diagnosis Related Group.

We also examined whether the differential adoption of hospital outpatient TKA between mandatory CJR and control hospitals explained the reduction in the complexity of the elective MS-DRG 470 episode group. If CJR hospitals retained TKA patients of lower complexity in the inpatient setting instead of providing the surgery in the outpatient setting, while control group hospitals shifted lower complexity patients to the outpatient setting, the average complexity of CJR participant hospitals' inpatient population could decrease relative to the control group. To test this, we performed a sensitivity analysis accounting for the differential outpatient TKA rates. We found that changes in the patient population not caused by the differential outpatient TKA rates resulted in a similar relative decrease in CJR episode payments (Appendix C). Thus the differential outpatient TKA rates do not explain these changes in the inpatient elective MS-DRG 470 patient population.

Analyses of patient characteristics. The analyses of changes in patient characteristics provide additional insight into the reduction in the complexity of CJR patients in the elective MS-DRG 470 episode group, relative to the control group. The results are generally consistent with those of the composite measure analyses. Specifically, the CJR elective MS-DRG 470 patient population became relatively less complex from the baseline to the intervention, with respect to eligibility for Medicaid, prior SNF utilization, and any prior care (Exhibit 32).

Medicaid. There was a statistically significant 1.3pp decrease in the proportion of patients who were eligible for Medicaid ($p<0.05$), a characteristic associated with higher expected episode payments.

Prior SNF stay. There was a 0.4pp decrease in the proportion of patients who had a SNF stay in the six months prior to their LEJR procedure ($p<0.01$).

Any prior care. There was a 1.0pp decrease in the proportion of patients who received any care in the six months prior to their LEJR procedure ($p<0.05$), where any care is defined as any inpatient hospital, psychiatric hospital, ED, IRF, SNF, HH, long-term care hospital (LTCH), or hospice utilization.

Exhibit 32: For the elective MS-DRG 470 episode group at mandatory CJR hospitals, three characteristics indicated a less complex patient population
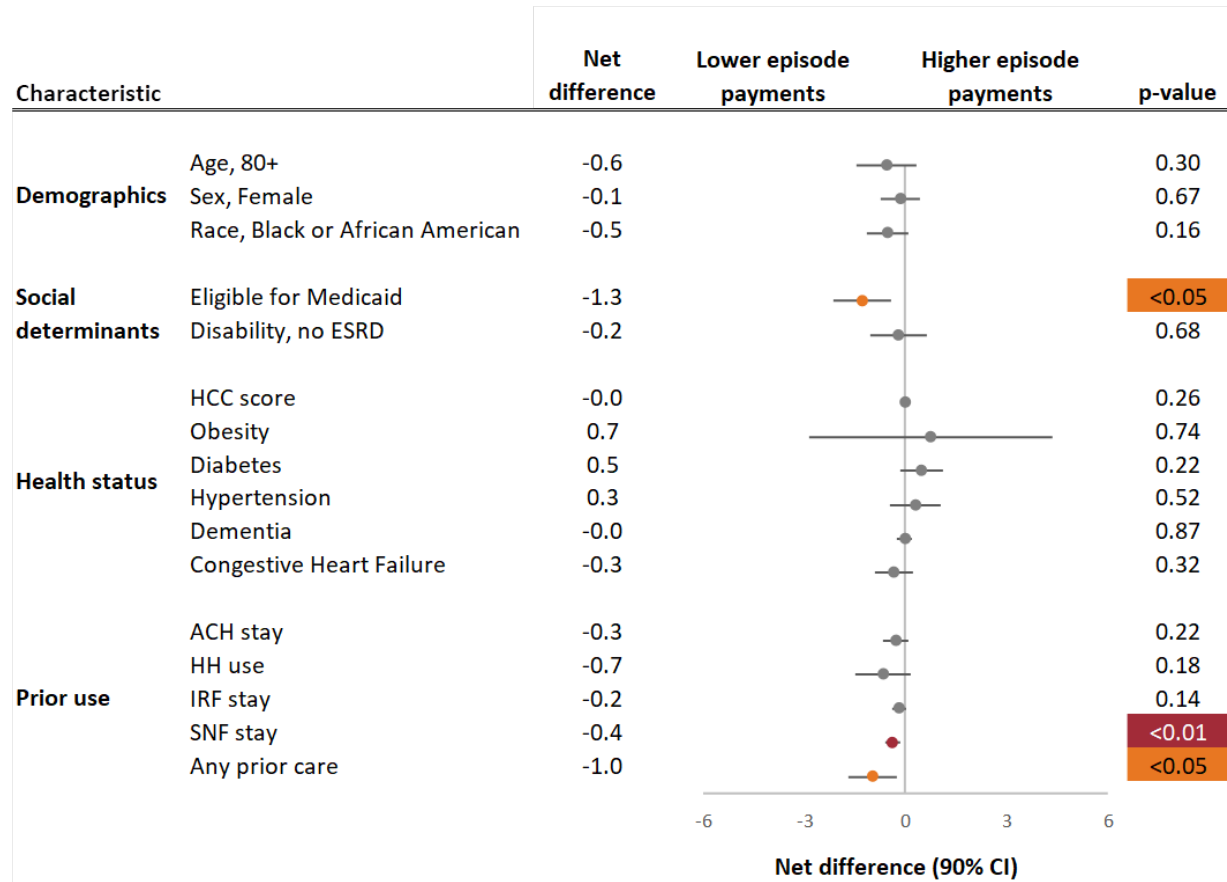
Characteristic		Net difference	Lower episode payments	Higher episode payments	p-value
Demographics	Age, 80+	-0.6			0.30
	Sex, Female	-0.1			0.67
	Race, Black or African American	-0.5			0.16
Social determinants	Eligible for Medicaid	-1.3			<0.05
	Disability, no ESRD	-0.2			0.68
Health status	HCC score	-0.0			0.26
	Obesity	0.7			0.74
	Diabetes	0.5			0.22
	Hypertension	0.3			0.52
	Dementia	-0.0			0.87
	Congestive Heart Failure	-0.3			0.32
Prior use	ACH stay	-0.3			0.22
	HH use	-0.7			0.18
	IRF stay	-0.2			0.14
	SNF stay	-0.4			<0.01
	Any prior care	-1.0			<0.05

Net difference (90% CI) [x-axis: -6, -3, 0, 3, 6]

Source: CJR evaluation team analysis of Medicare claims and enrollment data for episodes initiated in 2012 through 2014 that ended between April 2012 and March 2015 (baseline) and episodes initiated during or after April 2016 that ended by December 2019 (intervention).

Notes: Net differences that are significant at the 1%, 5%, or 10% significance level are indicated by red, orange, or yellow shading, respectively.

Any prior care includes inpatient hospital, psychiatric hospital, emergency department visits, IRF, SNF, HH, long-term care hospital, and hospice during the six months prior to anchor hospitalization.

ACH = acute care hospital, CI = confidence interval, ESRD = end-stage renal disease, HCC = hierarchical condition category, HH = home health, IRF = inpatient rehabilitation facility, MS-DRG = Medicare Severity-Diagnosis Related Group, SNF = skilled nursing facility.

Changes in patient characteristics for the other three episode groups did not indicate changes in patient complexity. For the elective MS-DRG 469 episode group, the changes in patient characteristics were not consistently associated with lower or higher episode payments (Appendix J). Likewise, for the fracture MS-DRG 470 and fracture MS-DRG 469 episode groups we also found no consistent evidence of relative changes in the mandatory CJR patient population (Appendix J).

c. Conclusion

After accounting for changes in LEJR patient complexity over time, elective MS-DRG 470 CJR patients were healthier in the intervention period (PY1-4) than the baseline period, relative to

control patients. We found similar results for mandatory CJR hospitals during PY1-3, which we discussed in an earlier report.[58]

It is difficult to determine whether this shift in patient mix is due to external factors outside of the hospital's control or intentional patient selection. Our analyses suggest that the relative decline in patient complexity does not appear to be due to changes in coding or LEJR volume. While the characteristics of patients receiving LEJR could change if hospitals better prepare patients for surgery, our analyses indicate there are relatively fewer CJR patients who are dual eligible, an indicator of lower socioeconomic status, or who had prior health care utilization. Changes in the proportion of patients with these characteristics are not consistent with optimizing patients prior to their surgery. Lastly, the differential outpatient TKA rates between CJR and control hospitals does not explain the decline in patient complexity.

Because target prices, and thus reconciliation payments, are not adjusted for patient complexity other than through fracture status (elective or fracture) and MS-DRG (470 or 469), the relative decline in patient complexity for the elective MS-DRG 470 episode group likely contributed to higher reconciliation payments. This is because lower patient complexity is associated with lower average payments which, in turn, are associated with increased reconciliation payments. As a result, at least a portion of the reconciliation payments to mandatory CJR hospitals were due to the relative decline in the complexity of their patients. In PY6 and beyond, target prices will be adjusted for additional characteristics, such as a beneficiary's dual eligibility status.[59] This will likely reduce the degree to which changes in patient complexity lead to changes in reconciliation payments because episode target prices may more accurately reflect differences in payments due to variation in patient complexity.

6c. What was the impact of the CJR model on payments in the 30 days following the episode?

We monitor payments for services provided after the episode because they might indicate that instead of care redesign, participants merely postponed services to reduce episode payments. While this would reduce payments during the episode, postponing services would not reduce overall Medicare spending.

a. Methods

The same DiD methods described in Section II.A.1.b were used for this analysis. Additional details about the methodology are included in Appendix C Section III.

Acronym	
DiD	difference-in-differences

---

[58] Centers for Medicare & Medicaid Services. Comprehensive care for joint replacement model - third annual report. https://innovation.cms.gov/data-and-reports/2020/cjr-thirdannrpt. 2020:98-106

[59] Centers for Medicare & Medicaid Services. Medicare Program: Comprehensive Care for Joint Replacement Model Three-Year Extension and Changes to Episode Definition and Pricing; Medicare and Medicaid Programs; Policies and Regulatory Revisions in Response to the COVID–19 Public Health Emergency; Final Rule 2021:1-81.

b.   Results

During the first four performance years, the CJR model had no impact on payments for services provided during the 30 days following the end of the episode (DiD = -$32, p=0.18, Appendix D, Exhibit D-1). However, we cannot be certain that there was no impact of the model on payments in the 30 days following the episode because this outcome failed parallel trends tests. Parallel trends is an assumption that underlies our methodological approach. Because failure to pass the parallel trends test indicates that CJR and control group hospitals may have had different patterns of post episode payments in the baseline, we cannot isolate the impact of the CJR model on this outcome. Additional details about the methodology are included in Appendix C Section III.

c.   Conclusion

The CJR model likely had no impact on payments for services provided in the 30 days following the end of the episode.

7.   What was the impact of the CJR model on health equity?

The estimates of the impact of the CJR model on payments, utilization, and quality discussed earlier in this report reflect the average experience across all patients with CJR episodes. Although those estimates indicate that the CJR model reduced average episode payments with no systematic evidence of reduced quality, the findings could differ for particular groups. Indeed, some have posited that the CJR model could exacerbate or reduce racial or socioeconomic disparities in access to LEJR or health outcomes.[60,61] For example, studies have found that Black or African American patients are more likely to be discharged to institutional PAC following their LEJR than white patients, which would likely result in costlier episodes on average.[62] If, for instance, CJR hospitals avoided Black or African American patients because of possible higher episode payments, the CJR model could widen existing disparities in access to LEJRs. Alternatively, the incentives for improving quality of care due to the CJR model's quality-adjusted target price could particularly benefit patient groups with historically poorer health outcomes and thus reduce disparities in care.[63]

To evaluate the impact of the CJR model on health equity, we studied subpopulations with historically poorer access to care and health outcomes for any differences in the impact of the CJR model on these subpopulations relative to the entire population under the model. We

---

[60]   Ibrahim SA, Kim H, McConnell KJ. The CMS Comprehensive Care Model and Racial Disparity in Joint Replacement. JAMA. 2016;316(12):1258–1259. doi:10.1001/jama.2016.12330

[61]   U.S. Department of Health and Human Services Office of the Assistant Secretary for Planning and Evaluation. (2016). *Report to Congress: Social Risk Factors and Performance Under Medicare's Value-Based Purchasing Programs*. https://aspe.hhs.gov/system/files/pdf/253971/ASPESESRTCfull.pdf

[62]   Singh JA, Kallan MJ, Chen Y, Parks ML, Ibrahim SA. Association of Race/Ethnicity With Hospital Discharge Disposition After Elective Total Knee Arthroplasty. JAMA Netw Open. 2019 Oct 2;2(10):e1914259. doi: 10.1001/jamanetworkopen.2019.14259. Erratum in: JAMA Netw Open. 2019 Dec 2;2(12):e1918528. PMID: 31664446; PMCID: PMC6824220.

[63]   Ibrahim SA, Kim H, McConnell KJ. The CMS Comprehensive Care Model and Racial Disparity in Joint Replacement. JAMA. 2016;316(12):1258–1259. doi:10.1001/jama.2016.12330

conducted claims- and survey-based analyses for three subpopulations: those who are Black or African American (Black/AA); those eligible for both Medicare and Medicaid (dually eligible or with dual eligibility), and those who are Black/AA with dual eligibility.

---

a. Key Findings

- The CJR model resulted in larger episode payment reductions for Black/AA patients than for white patients.

- Payment reductions for patients with dual eligibility and for Black/AA patients with dual eligibility were not statistically different from the payment reductions for non-dually eligible patients or white patients without dual eligibility, respectively.

- While our analyses provide limited evidence of a differential impact of the CJR model on patient subpopulations with historically poorer access to care and health outcomes, the results should be interpreted with caution given the possibility of changes in the population of patients receiving LEJR, which could inflate health outcomes without improving the quality of care.

---

b. Methods

Claims-based analysis

To understand the impact of the CJR model on health equity, we studied three subpopulations of patients—patients who are Black/AA, patients with dual eligibility for Medicare and Medicaid, and Black/AA patients with dual eligibility.[64] The comparison subpopulations are, respectively, white patients, patients who are not dually eligible, and white patients without dual eligibility. This analysis considered only patients at mandatory CJR hospitals.

Acronyms	
AA	African American
DiD	difference-in-differences
LEJR	lower extremity joint replacement
MSA	metropolitan statistical area
pp	percentage point
TKA	total knee arthroplasty

We used the same DiD approach used in our other analyses to estimate the impact of the CJR model on one of our three target patient populations and again on the corresponding comparison patient population. Then we estimated the difference between the two CJR impacts to determine the differential impact of the model on the target subpopulation. For example, we estimated the impact of the CJR model on patients with dual eligibility and estimated the impact on patients without dual eligibility. We then estimated the difference between the two estimated impacts to determine the differential impact of the CJR model on patients with dual eligibility. Our analysis uses the same propensity score weighting methodology as our other claims-based analyses.

---

[64] Findings for patients with a fracture, who may be medically vulnerable to changes in care, are in Section II.A.5.c.

It is important to note that our analysis implicitly includes only patients who received care. If CJR hospitals avoided patients who might require higher than average resources to treat, our analysis will attribute any effects of lower patient complexity on payments in the intervention period to the CJR model. Our analysis of changes in patient mix found evidence of patient mix changes under the CJR model for elective MS-DRG 470, but patient mix changes within a subpopulation may differ from changes across the entire patient population. We cannot be certain that differences in impacts for subpopulations are not due to changes in patient mix or volume. We will investigate any relationships between patient mix and subpopulation volume in a future report.

A full description of the methodology is in Appendix C (Section II and Section III).

Patient survey analysis

As described in Section II.A.5, we surveyed patients after the end of their inpatient LEJR episode on a variety of topics, including change in functional status and pain (recalled from before their surgery to the time of the survey), satisfaction with overall recovery, satisfaction with care management, experience with care transitions, and caregiver help needed after returning home. Patient survey data were collected in two waves that covered episodes with inpatient discharges from mandatory CJR hospitals in March or April 2019 and in September or October 2019.

For two subpopulations, Black/AA patients and patients with dual eligibility,[65] we estimated risk-adjusted differences between CJR and control respondents within each group and estimated the risk-adjusted differential impact between the target subpopulation and the comparison subpopulation. Mirroring the claims analysis, the comparison groups are white respondents and respondents without dual eligibility, respectively. In this analysis, we accounted for beneficiary, hospital, and MSA attributes. Please see Section II.A.5, Appendix C, and Appendix G for more detail on these methods.

c.   Results

Differential impact of the CJR model on total episode payments

For mandatory CJR hospitals, the CJR model resulted in relative reductions in average episode payments for each subpopulation and for each comparison group (Exhibit 33). The change in average episode payments ranged from a reduction of $1,374 (p<0.01) for white patients without dual eligibility to a reduction of $2,435 (p<0.01) for the Black/AA population. The relative reduction in average episode payments was greater for the Black/AA population than for the white population (-$1,031, p<0.05).

---

[65]   Due to sample size considerations, the Black/AA subpopulation with dual eligibility was not included in the analyses of the patient survey. See Appendix C (Exhibit C-15) for sample sizes and response rates.

Exhibit 33:   The CJR model resulted in a larger reduction in total episode payments for Black/AA patients relative to white patients

Subpopulation	Comparison population	Subpopulation DiD estimate (total payments)	Comparison population DiD estimate (total payments)	Differential CJR impact
Black/AA	White	-$2,435	-$1,404	-$1,031
Dually eligible	Not dually eligible	-$2,017	-$1,441	-$576
Black/AA with dual eligibility	White without dual eligibility	-$2,215	-$1,374	-$841

Source: CJR evaluation team analysis of Medicare claims and enrollment data for LEJR episodes initiated in 2012 through 2014 that ended between April 2012 and March 2015 (baseline) and LEJR episodes initiated during or after April 2016 that ended by December 2019 (intervention).

Notes: Estimates that are significant at the 1%, 5%, or 10% significance level are indicated by red, orange, or yellow shaded cells, respectively.

Because CJR participant hospitals shifted a lower share of TKAs to the hospital outpatient setting, the control group includes outpatient TKA episodes that have been weighted to balance the episode volume in the CJR hospitals.

AA = African American, DiD = difference-in-differences, LEJR = lower extremity joint replacement, TKA = total knee arthroplasty.

Differential impact of the CJR model on quality of care for Black/AA patients

The CJR model was associated with a 0.41 percentage point decrease in the all-cause mortality rate for Black/AA patients (p<0.05) and no statistically significant change in the all-cause mortality rate for white patients (Exhibit 34). The CJR model resulted in a larger reduction in the mortality rate for Black/AA patients than for white patients (-0.48 percentage points, p<0.05).

Exhibit 34:   The CJR model resulted in a larger reduction in all-cause mortality for Black/AA patients relative to white patients

Outcome	Black/AA DiD estimate (pp)	White DiD estimate (pp)	Differential CJR impact (pp)
Mortality	-0.41	0.06	-0.48
Emergency department use	-0.29	0.09	-0.37
Unplanned readmission rate	-0.50	-0.23	-0.28

Source: CJR evaluation team analysis of Medicare claims and enrollment data for LEJR episodes initiated in 2012 through 2014 that ended between April 2012 and March 2015 (baseline) and LEJR episodes initiated during or after April 2016 that ended by December 2019 (intervention).

Notes: All estimates in this exhibit reflect percentage points. Estimates that are significant at the 1%, 5%, or 10% significance level are indicated by red, orange, or yellow shaded cells, respectively.

Because CJR participant hospitals shifted a lower share of TKAs to the hospital outpatient setting, the control group includes outpatient TKA episodes that have been weighted to balance the episode volume in the CJR hospitals.

AA = African American, DiD = difference-in-differences, LEJR = lower extremity joint replacement, pp = percentage point, TKA = total knee arthroplasty.

The estimated effect of the CJR model on emergency department use and on the rate of unplanned readmissions was not statistically significant for either subpopulation, and the estimated differential impacts were likewise not statistically significant.

Another way of understanding the differential impact of the CJR model is to examine the change in the differences in an outcome between the subpopulation of interest and its comparison population (Exhibit 35). The all-cause mortality rate for Black/AA patients in mandatory hospitals was 1.2 percentage points lower than for white patients in mandatory hospitals during the baseline period (1.6% for Black/AA and 2.8% for white, Appendix N Exhibit N-1). This gap in mortality grew by 39.7% because the CJR model reduced the all-cause mortality rate for Black/AA patients by 0.48 percentage points more than for white patients. The differential impact is driven by two changes—the mortality gap grew between the baseline and intervention periods for patients at mandatory CJR hospitals, but narrowed for patients at control hospitals. Specifically, at mandatory CJR hospitals, the mortality rate for Black/AA patients decreased from 1.6% to 1.3% and the mortality rate for white patients decreased from 2.8% to 2.6% (Appendix N). At control hospitals, the mortality rate for Black/AA patients increased from 1.6% to 1.8% and the mortality rate for white patients decreased from 2.9% to 2.6% (Appendix N).

Exhibit 35: The favorable impact of the CJR model on mortality for Black/AA patients expanded the gap in mortality rates between Black/AA and white patients by nearly 40%

	Black/AA patients vs white patients		
Outcome	**Baseline outcome gap, risk adjusted – positive values indicate higher rates for Black/AA patients (pp)**	**Differential CJR impact between Black/AA and White patients (pp) [a]**	**Differential impact as a % of baseline gap**
Mortality	-1.2	-0.48	39.7%
Emergency department use	5.1	-0.37	-7.3%
Unplanned readmission rate	1.5	-0.28	-18.6%

Source: CJR evaluation team analysis of Medicare claims and enrollment data for LEJR episodes initiated in 2012 through 2014 that ended between April 2012 and March 2015 (baseline) and LEJR episodes initiated during or after April 2016 that ended by December 2019 (intervention).

Notes: The baseline outcome gap is the average outcome for Black/AA patients at mandatory CJR hospitals during the baseline minus the average outcome for white patients at mandatory CJR hospitals during the baseline.

Because CJR participant hospitals shifted a lower share of TKAs to the hospital outpatient setting, the control group includes outpatient TKA episodes that have been weighted to balance the episode volume in the CJR hospitals.

AA = African American, LEJR = lower extremity joint replacement, pp = percentage point, TKA = total knee arthroplasty.

[a] Estimates that are significant at the 1%, 5%, or 10% significance level are indicated by red, orange, or yellow shaded cells, respectively.

The patient survey did not reveal any differential impacts of the CJR model on changes in functional status for Black/AA respondents relative to white respondents (Exhibit 36).[66] The mean changes in self-reported functional status were similar between CJR and control for both Black/AA and white respondents.

Exhibit 36: There were similar changes in functional status for Black/AA and white survey respondents under the CJR model

Survey measure	Response range of survey measure	Difference of mean change in self-reported measure between CJR and control[a] (higher values indicate a more favorable change)		Differential impact of the CJR model
		Black/AA	**White**	
Ability to walk by yourself without resting	-4 to 4	0.01	0.06	-0.04
Difficulty walking up or down 12 stairs	-3 to 3	0.04	-0.01	0.05
Difficulty rising from sitting	-4 to 4	0.03	0.00	0.02
Difficulty standing	-4 to 4	0.09	-0.01	0.10
Use of a mobility aid	-2 to 2	0.03	0.01	0.02
Difficulty getting on/off the toilet	-4 to 4	0.05	-0.01	0.06
Frequency that pain interferes with normal activities	-4 to 4	0.06	-0.01	0.07
Medication use for pain in the joint you had replaced	-3 to 3	-0.03	0.02	-0.05

Source: CJR evaluation team analysis of patient survey data for episodes with discharge in March, April, September, or October 2019.

Notes: The estimates in this exhibit are the results of a cross-sectional regression model, weighted for sampling and nonresponse. Estimates that are significant at the 1%, 5%, or 10% significance level are indicated by red, orange, or yellow shaded cells, respectively.

AA = African American

a The change in a given measure refers to the difference between a respondent's self-reported status at the time of the survey and the respondent's recalled status prior to the hospitalization. Higher values represent a more favorable change.

Survey measures on satisfaction with overall recovery, experience with care transitions, and caregiver help all indicated no differential impact between Black/AA and white respondents (Appendix N, Exhibits N-8, N-9, and N-10).

The CJR model did have a differential impact on satisfaction with care management. For Black/AA respondents, the CJR patients had higher satisfaction with their care management than control beneficiaries in regard to health care providers listening to their preferences (p<0.01; Exhibit 37). Comparing this result to that of white respondents, we see the CJR model had a

---

[66] Patient survey measures have been constructed so that higher values represent a more favorable change.

differential impact of 6.91 survey measure points (p<0.01) for Black/AA respondents. This result contributed to a positive differential impact in the composite measure of satisfaction with care management (p<0.1).

Exhibit 37:   Under the CJR model Black/AA respondents were more satisfied than white respondents in the extent to which providers listened to their preferences

Survey measure	Response range of survey measure	Difference between CJR and control		Differential impact of the CJR model
		Black/AA	**White**	
Composite measure of satisfaction with care management	0 to 100	2.87	-0.32	3.18
Health care providers listened to preferences	0 to 100	6.26	-0.65	6.91
Satisfaction with discharge destination	0 to 100	1.39	0.18	1.22
Satisfaction with care coordination	0 to 100	1.07	-0.45	1.52
Satisfaction with treatment instructions	0 to 100	2.46	-0.52	2.98

Source: CJR evaluation team analysis of patient survey data for episodes with discharge in March, April, September, or October 2019.

Notes: The estimates in this exhibit are the results of a cross-sectional regression model, weighted for sampling and nonresponse. Estimates that are significant at the 1%, 5%, or 10% significance level are indicated by red, orange, or yellow shaded cells, respectively.

AA = African American.

Measures of satisfaction with care management are scaled from 0 to 100 points, where 0 = "very dissatisfied," 25 = "somewhat satisfied," 50 = "neither satisfied nor dissatisfied," 75 = "somewhat satisfied," and 100 = "very satisfied"

Differential impact of the CJR model on quality of care for patients with dual eligibility

For patients with dual eligibility, the CJR model was associated with a 0.73 percentage point decrease in the rate of unplanned readmissions (p<0.10), but this impact was not statistically different from the impact of the CJR model on patients without dual eligibility (Exhibit 38). No other estimated impact is statistically significant for patients with dual eligibility compared to patients without dual eligibility.

Exhibit 38: The CJR model did not result in differential quality impacts for patients with dual eligibility

Outcome	Patients with dual eligibility DiD estimate (pp)	Patients without dual eligibility DiD estimate (pp)	Differential CJR impact (pp)
Mortality	0.31	-0.02	0.33
Emergency department use	0.00	0.14	-0.14
Unplanned readmission rate	-0.73	-0.25	-0.48

Source: CJR evaluation team analysis of Medicare claims and enrollment data for LEJR episodes initiated in 2012 through 2014 that ended between April 2012 and March 2015 (baseline) and LEJR episodes initiated during or after April 2016 that ended by December 2019 (intervention).

Notes: All estimates in this exhibit reflect percentage points. Estimates that are significant at the 1%, 5%, or 10% significance level are indicated by red, orange, or yellow shaded cells, respectively.

Because CJR participant hospitals shifted a lower share of TKAs to the hospital outpatient setting, the control group includes outpatient TKA episodes that have been weighted to balance the episode volume in the CJR hospitals.

DiD = difference-in-differences, LEJR = lower extremity joint replacement, pp = percentage point, TKA = total knee arthroplasty.

Although none of the estimated differential impacts are statistically significant for patients with dual eligibility, our estimate indicates the CJR model increased mortality for patients with dual eligibility, relative to patients without dual eligibility (Exhibit 39). Patients with dual eligibility had a 1.3 percentage point higher mortality rate in the baseline period than patients without dual eligibility, and this gap expanded by nearly 25% under the model, although this increase was not statistically significant. This differential impact is driven primarily by a larger decrease in mortality for dually eligible control patients (who had a higher baseline mortality rate) than for dually eligible patients at mandatory CJR hospitals. The mortality rate for patients with dual eligibility at control hospitals decreased from 4.3% in the baseline to 3.9% in the intervention period. By contrast, the mortality rate for patients with dual eligibility at mandatory CJR hospitals decreased from 3.8% in the baseline to 3.7% in the intervention.

Exhibit 39: There were no differential quality impacts of the CJR model on patients with and without dual eligibility

	Patients with dual eligibility vs patients without dual eligibility		
Outcome	**Baseline outcome gap, risk adjusted – positive values indicate higher rates for patients with dual eligibility (pp)**	**Differential CJR impact between patients with dual eligibility and patients without dual eligibility (pp)[a]**	**Differential impact as a % of baseline gap**
Mortality	1.3	0.33	24.7%
Emergency department use	6.9	-0.14	-2.0%
Unplanned readmission rate	5.0	-0.48	-9.6%

Source: CJR evaluation team analysis of Medicare claims and enrollment data for LEJR episodes initiated in 2012 through 2014 that ended between April 2012 and March 2015 (baseline) and LEJR episodes initiated during or after April 2016 that ended by December 2019 (intervention).

Notes: The baseline outcome gap is defined as the average outcome for patients with dual eligibility at mandatory CJR hospitals minus the average outcome value for patients without dual eligibility at mandatory CJR hospitals.

   Because CJR participant hospitals shifted a lower share of TKAs to the hospital outpatient setting, the control group includes outpatient TKA episodes that have been weighted to balance the episode volume in the CJR hospitals.

   LEJR = lower extremity joint replacement, pp = percentage point, TKA = total knee arthroplasty.

   [a] Estimates that are significant at the 1%, 5%, or 10% significance level are indicated by red, orange, or yellow shaded cells, respectively.

Based on the patient survey, most of the mean changes in self-reported functional status were similar between CJR and control respondents with and without dual eligibility (Exhibit 40). For those with dual eligibility, CJR respondents reported having more improvement walking up or down stairs than control respondents ($p<0.1$). Among the subpopulation without dual eligibility, CJR respondents reported having more improvement in walking than control respondents ($p<0.05$). Comparing the impacts between the subpopulations, the subpopulation with dual eligibility experienced more improvement under the CJR model in walking up or down stairs ($p<0.05$) relative to the control respondents without dual eligibility.

Exhibit 40:   The CJR model did not have systematically differential impacts on improvements in functional status for respondents with dual eligibility

| Survey measure | Response range | Difference of mean change in self-reported measure in between CJR and control[a] | | Differential impact of the CJR model |
		With dual eligibility	Without dual eligibility	
Ability to walk by yourself without resting	-4 to 4	0.04	0.04	-0.01
Difficulty walking up or down 12 stairs	-3 to 3	0.09	-0.02	0.11
Difficulty rising from sitting	-4 to 4	0.06	0.00	0.07
Difficulty standing	-4 to 4	0.04	-0.00	0.04
Use of a mobility aid	-2 to 2	-0.01	0.01	-0.01
Difficulty getting on/off the toilet	-4 to 4	0.03	-0.01	0.04
Frequency that pain interferes with normal activities	-4 to 4	0.06	0.00	0.06
Medication use for pain in the joint you had replaced	-3 to 3	0.01	0.02	-0.01

Source: CJR evaluation team analysis of patient survey data for episodes with discharge in March, April, September, or October 2019.

Notes: The estimates in this exhibit are the results of a cross-sectional regression model, weighted for sampling and nonresponse. Estimates that are significant at the 1%, 5%, or 10% significance level are indicated by red, orange, or yellow shaded cells, respectively.

[a]   The change in a given measure refers to the difference between a respondent's self-reported status at the time of the survey and the respondent's recalled status prior to the hospitalization. Higher values represent a more favorable change.

Survey measures on satisfaction with overall recovery, satisfaction with care management, experience with care transitions, and caregiver help all indicated no differential impact of the CJR model between survey respondents with dual eligibility and those without (Appendix N, Exhibit N-12, N-13, and N-14).

Differential impact of the CJR model on Black/AA patients with dual eligibility

There was no differential impact of the CJR model for Black/AA patients with dual eligibility compared with white patients without dual eligibility (Exhibit 41). This subpopulation is substantially smaller than the others (more than half of Black/AA patients in our sample are not dually eligible). Because of the smaller sample size, differences would have to be much larger to detect any differences in impact of the CJR model between the two subpopulations, so it is possible that there were differential impacts, despite the lack of statistically significant results.

Exhibit 41:   The CJR model did not result in any differential impacts on quality for Black/AA patients with dual eligibility or white patients without dual eligibility

Outcome	Black/AA patients with dual eligibility DiD estimate (pp)	White patients without dual eligibility DiD estimate (pp)	Differential CJR impact (pp)
Mortality	-0.61	-0.00	-0.61
Emergency department use	-1.49	0.15	-1.64
Unplanned readmission rate	0.03	-0.18	0.21

Source: CJR evaluation team analysis of Medicare claims and enrollment data for LEJR episodes initiated in 2012 through 2014 that ended between April 2012 and March 2015 (baseline) and LEJR episodes initiated during or after April 2016 that ended by December 2019 (intervention).

Notes: All estimates in this exhibit reflect percentage points. Estimates that are significant at the 1%, 5%, or 10% significance level are indicated by red, orange, or yellow shaded cells, respectively.

Because CJR participant hospitals shifted a lower share of TKAs to the hospital outpatient setting, the control group includes outpatient TKA episodes that have been weighted to balance the episode volume in the CJR hospitals.

AA = African American, DiD = difference-in-differences, LEJR = lower extremity joint replacement, pp = percentage point, TKA = total knee arthroplasty.

Exhibit 42: **The CJR model did not result in any differential impacts on quality for Black/AA patients with dual eligibility compared to white patients without dual eligibility**

Outcome	Black/AA patients with dual eligibility vs white patients without dual eligibility		
	Baseline outcome gap, risk adjusted – positive values indicate higher rates for Black/AA patients with dual eligibility (pp)	**Differential CJR impact between Black/AA patients with dual eligibility and white patients without dual eligibility (pp)** [a]	**Differential impact as a % of baseline gap**
Mortality	-0.6	-0.61	103.1%
Emergency department use	11.6	-1.64	-14.1%
Unplanned readmission rate	5.2	0.21	4.0%

Source: CJR evaluation team analysis of Medicare claims and enrollment data for LEJR episodes initiated in 2012 through 2014 that ended between April 2012 and March 2015 (baseline) and LEJR episodes initiated during or after April 2016 that ended by December 2019 (intervention).

Notes: The baseline outcome gap is defined as the average outcome for Black/AA patients with and dual eligibility at mandatory CJR hospitals minus the average outcome value for white patients without dual eligibility at mandatory CJR hospitals.

Because CJR participant hospitals shifted a lower share of TKAs to the hospital outpatient setting, the control group includes outpatient TKA episodes that have been weighted to balance the episode volume in the CJR hospitals.

AA = African American, LEJR = lower extremity joint replacement, pp = percentage point, TKA = total knee arthroplasty.

[a] Estimates that are significant at the 1%, 5%, or 10% significance level are indicated by red, orange, or yellow shaded cells, respectively.

The CJR model did not have a significantly different impact on any of the three quality measures for Black/AA patients with dual eligibility relative to white patients without dual eligibility (Exhibit 42).

d. Conclusion

In both our claims- and survey-based analyses, we found limited evidence of different impacts of the CJR model on patient subpopulations with historically worse access to care and health outcomes relative to their comparison subpopulations. For most quality outcomes, estimated differential impacts were not statistically significant. For a few outcomes that did indicate a statistically significant differential impact, we found that the CJR model improved the quality of care for the subpopulation of interest. To fully understand the impact of the CJR model on health equity, however, it is necessary to evaluate whether the CJR model resulted in changes in access to care for certain subpopulations, which could improve health outcome measures by exacerbating disparities in access. Therefore, we urge caution in interpreting these results without concurrent results on patient access.

Differential changes in LEJR volume for a subpopulation could explain certain results in this analysis, like the decrease in mortality rate for Black/AA patients. In previous annual reports, our analysis of patient mix indicated that CJR hospitals had a lower proportion of Black/AA patients than control hospitals.[67,68] Other research has found that the CJR model is associated with a decrease in the likelihood that Black/AA patients received elective LEJRs.[69] If the rate of elective LEJRs for Black/AA patients has declined under the CJR model, with a resulting decline in patient mix for Black/AA patients who do receive LEJRs, then the relative improvements in quality under the CJR model for Black/AA patients could be due to changes in the patient population rather than actual improvements in the quality of care. Although we did not find relative decreases in Black/AA patients under the CJR model in the analyses in this report, changes in volume and resulting changes in patient mix could have unintended consequences that we have not detected. In future reports, we will examine the impact of the CJR model on LEJR volume and patient mix in patient subpopulations with historically poorer access to care and health outcomes.

8.　*What was the impact of the CJR model on hospitals in voluntary MSAs?*

At the start of PY3, the number of mandatory MSAs in the CJR model was scaled back from the 67 original randomly selected MSAs to the 34 MSAs with the highest average historical payments. While other sections of this report focus on CJR hospitals in the 34 *mandatory* MSAs that were continuously required to participate throughout the entire model, this section focuses on CJR hospitals in the 33 *voluntary* MSAs. In January 2018, CJR hospitals in these voluntary MSAs were given a one-time opportunity to opt-in to the CJR model for PY3-5. We therefore classify these voluntary CJR hospitals into two groups: 74 *opt-in* CJR hospitals, hospitals that continued their participation, and 200 *non-opt-in* CJR hospitals, hospitals that ended their participation in the CJR model.[70]

In the third annual evaluation report, we analyzed the characteristics and reconciliation amounts of voluntary CJR hospitals.[71] In this section, we extend this analysis by looking at the impact of the model on these CJR hospitals. To account for the ability of opt-in and non-opt-in hospitals to select to continue in the model, we constructed two comparison groups, one for opt-in hospitals and one for non-opt-in hospitals. The comparison groups are designed to be similar to

---

[67] Centers for Medicare & Medicaid Services. Comprehensive care for joint replacement model - second annual report. https://innovation.cms.gov/files/reports/cjr-secondannrpt.pdf. 2020: 52-57.

[68] Centers for Medicare & Medicaid Services. Comprehensive care for joint replacement model - third annual report. https://innovation.cms.gov/data-and-reports/2020/cjr-thirdannrpt. 2020: 98-106.

[69] Kim et al. (2021) Association of Medicare Mandatory Bundled Payment Program with the Receipt of Elective Hip and Knee Replacement in White, Black, and Hispanic Beneficiaries, *JAMA Network Open* 2021;4(3): e211772-e211772.

[70] These counts of hospitals refer to the voluntary CJR hospitals included in the analysis presented in this chapter. Please see Appendix C (Section II.C) for additional details about this methodology.

[71] Centers for Medicare & Medicaid Services. Comprehensive care for joint replacement model - third annual report. https://innovation.cms.gov/data-and-reports/2020/cjr-thirdannrpt. 2020:21-31.

the CJR hospitals they represent with respect to baseline characteristics that could affect a hospital's decision to opt-in or not opt-in.

The opt-in and non-opt-in CJR groups studied in this section include only CJR hospitals in the 33 *voluntary* MSAs. They do not include CJR hospitals located in mandatory MSAs that had a choice to voluntarily continue in the CJR model due to being designated as low-volume or rural. We do not include the low-volume and rural hospitals located in the 34 mandatory MSAs in the analysis of opt-in and non-opt-in hospitals in voluntary MSAs because low-volume and rural hospitals differ in important ways that would make them less comparable to the voluntary control group. First, by construction, the low-volume and rural hospitals are located in MSAs with higher average historical payments. Second, low-volume and rural hospitals are located in MSAs in which most hospitals are participating in CJR. Since an analysis of low-volume and rural hospitals would also need to account for their ability to select to continue in the model, we do not include them in the analysis of the CJR hospitals in the 34 *mandatory* MSAs that were continuously required to participate throughout the entire model.

a. Key Findings

- For opt-in CJR hospitals:

    - Average episode payments decreased by $752 more than for the matched control group during the first four performance years, equating to a 3.2% decrease from the baseline.

    - The reduction in average episode payments was mainly driven by a relative reduction in SNF payments due to a relative reduction in the average length of SNF stays.

    - There was a relative reduction in ED use of approximately 5.9% during the first four performance years. However, patients discharged to an IRF, SNF, or HHA had less improvement in functional status while in the PAC setting than patients in the matched control group.

- For non-opt-in CJR hospitals:

    - When non-opt-in CJR hospitals participated in the model (PY1-2), average episode payments decreased by $440 relative to the matched control group hospitals (1.7% decrease from the baseline). After they stopped participating (PY3-4), we did not find evidence that the reduction in average episode payments persisted.

    - During PY1-2, there were relative reductions in the average length of SNF stays and number of HHA visits among SNF and HHA users, respectively. These results persisted during PY3-4.

b. Methods

This analysis used a DiD design similar to the one used throughout this report.

One notable difference in this methodology from the methodology used for mandatory CJR hospitals is the construction of matched control groups. Hospitals in voluntary CJR MSAs chose to continue in the model, and thus CJR hospitals selected into the opt-in or non-opt-in hospital groups. This selection was correlated with many factors and characteristics that may have affected a given hospital's performance under the CJR model.[72] To account for this endogenous selection, we constructed subsets of the control group hospitals to use

Acronyms	
APM	alternative payment model
DiD	difference-in-differences
ED	emergency department
HH	home health
HHA	home health agency
IRF	inpatient rehabilitation facility
IRF-PAI	inpatient rehabilitation facility-patient assessment instrument
LEJR	lower extremity joint replacement
MDS	minimum data set
MSA	metropolitan statistical area
OASIS	outcome and assessment information set
OT	occupational therapy
PAC	post-acute care
PT	physical therapy
pp	percentage point
PY	performance year
SNF	skilled nursing facility
TKA	total knee arthroplasty

as a counterfactual when evaluating the impact of the voluntary CJR hospitals. Using propensity score matching, we selected control group hospitals located in voluntary control MSAs that resembled the voluntary CJR hospitals across a variety of baseline characteristics. We preformed separate matching procedures for opt-in CJR hospitals and non-opt-in CJR hospitals to create separate control group hospitals for the corresponding DiD and descriptive analyses. This approach uses the original randomized controlled trial design of the CJR model, while also modeling the decisions of the voluntary CJR hospitals about whether to continue in the CJR model.

For most analyses pertaining to *opt-in* CJR hospitals, we account for differential rates of outpatient TKAs between CJR and control hospitals. To do so, we include all matched control outpatient TKAs and weight them to represent the proportion of opt-in CJR inpatient TKAs that are predicted to have been outpatient TKAs in the absence of the CJR model. This approach provides an unbiased estimate of the impact of the CJR model on inpatient episodes for opt-in hospitals. The one exception is the analyses of PAC assessment measures, for which we used only LEJRs performed in the inpatient setting, because Medicare does not cover SNF care following outpatient procedures and outpatient TKA patients being first discharged to an

---

[72] Centers for Medicare & Medicaid Services. Comprehensive care for joint replacement model - third annual report. https://innovation.cms.gov/data-and-reports/2020/cjr-thirdannrpt. 2020:28-31.

IRF are rare. (Please see Section II.A.1.b and Appendix C, Section II for additional details on this method.)

All analyses pertaining to *non-opt-in* CJR hospitals included only LEJRs performed in the inpatient setting because Medicare coverage of outpatient TKAs did not begin until after non-opt-in hospitals stopped participating in the CJR model.
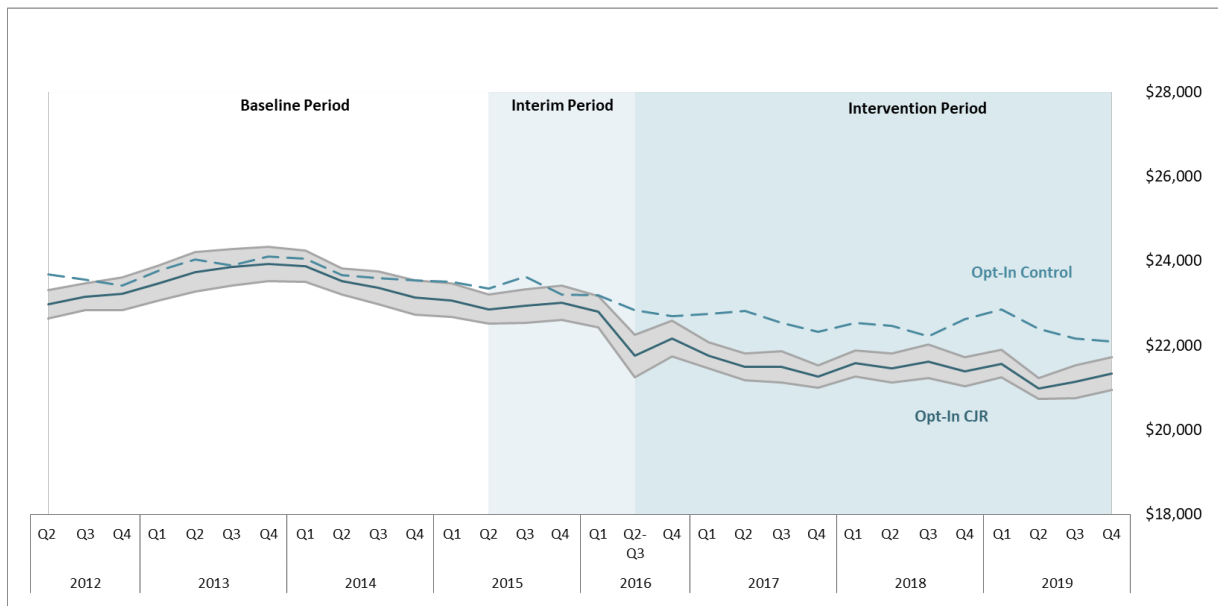
Additional details about the methodology employed in this section are available in Appendix C (Section II and Section III).

c. Results

Opt-in CJR hospitals in voluntary MSAs

During the first four performance years, average episode payments declined for both opt-in CJR hospitals and the matched control group hospitals, but decreased more for episodes initiated at CJR hospitals (Exhibit 43). Opt-in CJR hospitals had a $752 relative reduction in average episode payments, or approximately a 3.2% decrease from their baseline per-episode average payment (p<0.01; Appendix D, Exhibit D-3).

Exhibit 43: Average episode payments declined more for opt-in CJR hospitals than matched control hospitals in PY1-4

Source: CJR evaluation team analysis of Medicare claims and enrollment data for episodes initiated on or after January 2012 that ended by December 2019.

Notes: Episodes that ended between April 1, 2015 and March 31, 2016 (the interim period) were excluded from our baseline because the CJR model was announced in July 2015 and hospitals may have been preparing for their future participation in the CJR model during that time.

Because CJR participant hospitals shifted a lower share of TKAs to the hospital outpatient setting, the control group includes outpatient TKA episodes that have been weighted to balance the episode volume in the CJR hospitals.

The gray shading represents the 90% confidence interval for the CJR estimate.

PY = performance year, TKA = total knee arthroplasty.

During each of the first four performance years, opt-in CJR hospitals had a relative reduction in average episode payments, with the smallest per-episode reduction in PY1 (-$520, p<0.1) and the largest per-episode reduction in PY4 (-$985, p<0.01) (Exhibit 44).

Exhibit 44: Average episode payments for opt-in CJR hospitals decreased, with the largest decrease in PY4

Source: CJR evaluation team analysis of Medicare claims and enrollment data for episodes initiated in 2012 through 2014 that ended between April 2012 and March 2015 (baseline) and episodes initiated during or after April 2016 that ended by December 2019 (intervention).

Notes: The estimates in this exhibit are the result of a DiD model. DiD estimates that are significant at the 1%, 5%, or 10% significance level are indicated by red, orange, or yellow shaded circles, respectively. The whiskers represent 90% confidence intervals.
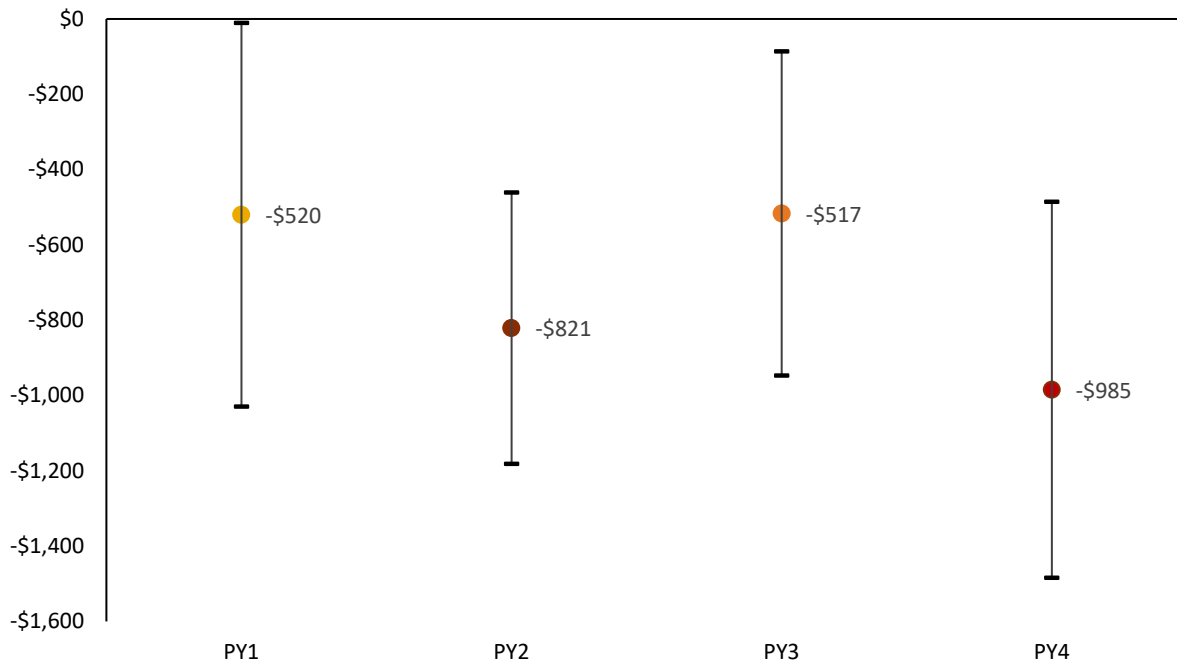
Because CJR participant hospitals shifted a lower share of TKAs to the hospital outpatient setting, the control group includes outpatient TKA episodes that have been weighted to balance the episode volume in the CJR hospitals.

PY = performance year, TKA = total knee arthroplasty.

The reduction in episode payments was mainly driven by a relative reduction in SNF payments. Average SNF payments decreased by $456 more for episodes initiated at opt-in CJR hospitals than for episodes initiated at matched control hospitals (10.5% from the CJR baseline, p<0.05; Appendix D, Exhibit D-3).

In addition, we also observed a relative reduction in intensity of SNF and HH care use. For patients with at least one stay, the average number of SNF days decreased by 2.8 days more for CJR episodes than for control group episodes from the baseline to the intervention period (12.2% from the CJR baseline, p<0.01; Appendix D, Exhibit D-3). For patients with at least one HH

visit, the average number of HH visits decreased by 0.9 visits more for CJR episodes than for control group episodes (6.6% from the CJR baseline, p<0.05). However, the CJR model did not have a statistically significant impact on the proportion of patients discharged to an IRF, SNF, or HHA, indicating that opt-in CJR hospitals may not have shifted from more intensive to less intensive PAC settings in response to the CJR model.

Claims-based measures of quality of care improved or did not change for opt-in hospitals under the CJR model. There was a relative reduction in the 90-day ED use rate. The 90-day ED use rate increased less from the baseline to the intervention period for opt-in CJR hospitals than for the control group (13.6% to 13.8% and 13.3% to 14.3%, respectively), resulting in a 0.8pp relative reduction for CJR hospitals (5.9% decrease from the CJR baseline, p<0.05; Appendix D, Exhibit D-3). The CJR model had no statistically significant impact on the unplanned readmission rate, mortality rate, or elective LEJR-specific complication rate.

Although claims-based analyses indicated quality of care was improved or maintained, we observed relative decreases in functional status in all three PAC settings. For CJR patients first discharged to an IRF, the average change in mobility score decreased by 1.3 points relative to the matched control group, or 12.0% from the CJR baseline (p<0.01, Exhibit 45). The average change in the mobility score remained the same for CJR patients (10.4 in both the baseline and the intervention), while the average change in the mobility score increased by 1.3 points for the control group over the same period (from 10.0 to 11.3) (Appendix D, Exhibit D-4).

Exhibit 45: Patients discharged from opt-in CJR hospitals had less improvement in functional status than control patients

First PAC discharge setting	Measure	DiD	DiD % of baseline	p-value
IRF	IRF mobility score	-1.3	-12.0%	p<0.01
SNF	Improved mobility	-5.4	-7.4%	p<0.01
	Improved toilet use	-5.3	-10.3%	p<0.01
	Without moderate to severe pain[a]	-1.0	-1.9%	p=0.68
HHA	Improved ambulation/locomotion	-0.9	-1.0%	p=0.22
	Improved bed transferring	-2.2	-2.6%	p<0.05
	Reduced pain	0.2	0.3%	p=0.88

Source: CJR evaluation team analysis of Medicare claims and enrollment data, IRF-PAI data, MDS data, and OASIS data for episodes initiated in 2012 through 2014 that ended between April 2012 and March 2015 (baseline) and episodes initiated during or after April 2016 that ended by September 2019 (intervention).

Notes: The estimates in this exhibit are the result of a DiD model. DiD estimates that are significant at the 1%, 5%, or 10% significance level are indicated by red, orange, or yellow shaded cells, respectively.

The relative change from CJR baseline is calculated as the DiD estimate as a percent of the CJR baseline level.

DiD = difference-in-differences, HHA = home health agency, IRF = inpatient rehabilitation facility, IRF-PAI = Inpatient Rehabilitation Facility Patient Assessment Instrument, MDS = Minimum Data Set, OASIS = Outcome and Assessment Information Set, PAC = post-acute care, SNF = skilled nursing facility.

[a] The pain measure for those initially discharged to a SNF was not risk adjusted following the specifications of the MDS 3.0 Quality Measure for short-stay patients used in the CMS Nursing Home Five-Star Rating System.

Among patients first discharged to a SNF, the proportion of CJR patients whose mobility improved during their SNF stay decreased by 5.4pp relative to the control group, or 7.4% from the CJR baseline (p<0.01). In addition, the proportion of CJR patients with improved toilet use decreased by 5.3pp relative to the control patients, or 10.3% from the CJR baseline (p<0.01). For both of these measures, the proportion of CJR patients with improved use decreased (for improved mobility, 73.1% in the baseline period to 69.8% in the intervention period; for improved toilet use, 51.0% in the baseline period to 45.9% in the intervention), while the proportion of control patients with improved use either remained constant or increased (Appendix D, Exhibit D-4). The direction of these changes could be an indication of a reduction of care for CJR patients during SNF stays.

For patients first discharged to an HHA, one out of three measures indicated a statistically significant relative decrease in functional status. The proportion of patients who improved in bed transferring decreased by 2.2pp relative to the control group, or 2.6% from the CJR baseline (p<0.05).

Non-opt-in CJR hospitals in voluntary MSAs

Similar to other hospital groups, average payments declined for both non-opt-in CJR hospitals and the matched control group hospitals following the baseline period (Exhibit 46). Average payments decreased by $440 more for non-opt-in CJR hospitals during the period when they were participating in the model (PY1-2) (p<0.05; Appendix D, Exhibit D-5 and Exhibit D-6). This relative reduction equates to a 1.7% decrease from the CJR baseline. After non-opt-in CJR hospitals stopped participating in the CJR model (PY3-4), there was a $286 relative reduction in average episode payments, however it was not statistically significant. Considering individual performance years, non-opt-in CJR hospitals had a statistically significant relative reduction in average episode payments during PY2 ($583 relative reduction per episode, p<0.01; Exhibit 47). The changes in payments in other performance years (PY1, PY3, and PY4) were smaller, statistically insignificant, and all of similar magnitude, ranging from a relative reduction of $133 to $372.

Exhibit 46: Non-opt-in CJR hospitals had a relative decline in average episode payments while participating in the model (PY1-2)
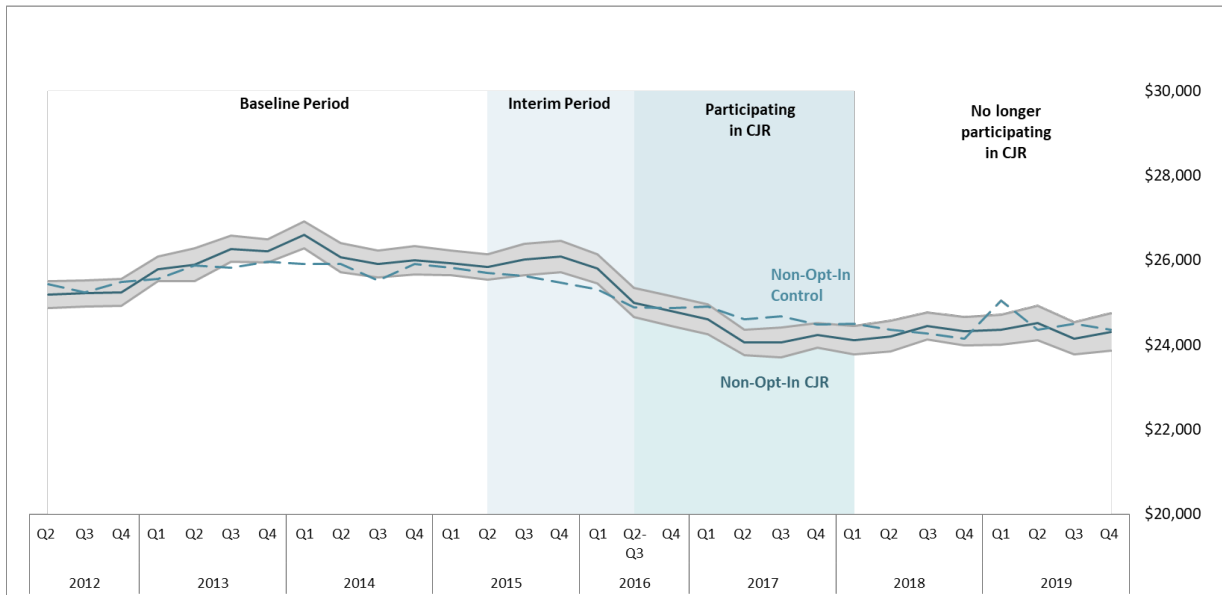
Source: CJR evaluation team analysis of Medicare claims and enrollment data for episodes initiated on or after January 2012 that ended by December 2019.

Notes: Episodes that ended between April 1, 2015 and March 31, 2016 (the interim period) were excluded from our baseline because the CJR model was announced in July 2015 and hospitals may have been preparing for their future participation in the CJR model during that time.

The gray shading represents the 90% confidence interval for the CJR estimate.

PY = performance year.

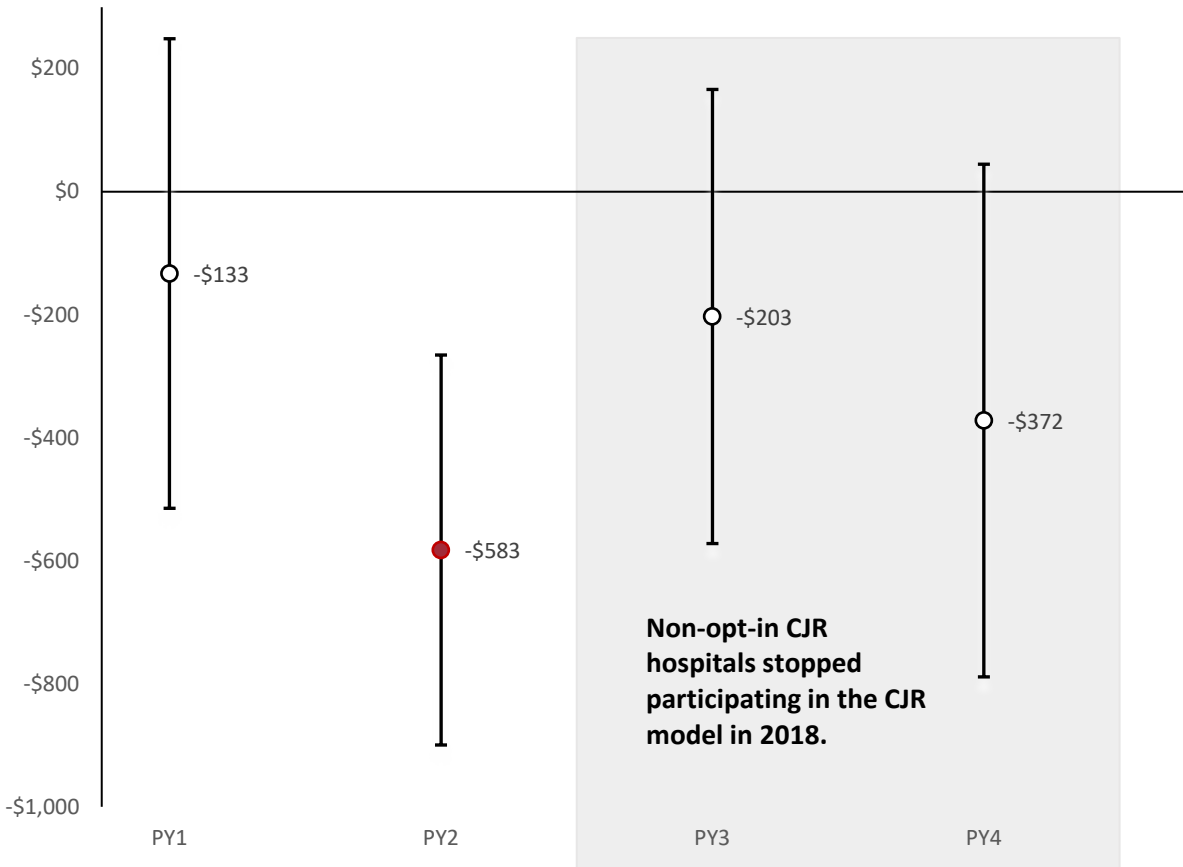**Exhibit 47: Non-opt-in CJR hospitals had a relative decline in average episode payments in PY2**

Source: CJR evaluation team analysis of Medicare claims and enrollment data for episodes initiated in 2012 through 2014 that ended between April 2012 and March 2015 (baseline) and episodes initiated during or after April 2016 that ended by December 2019.

Notes: The estimates in this exhibit are the result of a DiD model. DiD estimates that are significant at the 1%, 5%, or 10% significance level are indicated by red, orange, or yellow shaded circles, respectively. The whiskers represent 90% confidence intervals.

PY = performance year.

Moreover, analyses of service-level payments and service use suggest that after non-opt-in CJR hospitals stopped participating in the model, they may have shifted care from lower payment PAC settings (HHA) to higher payment PAC settings (SNF). Comparing PY3-4 to the baseline period, non-opt-in hospitals discharged a smaller proportion of patients to an HHA (5.3pp, p<0.05) and a larger proportion to a SNF (2.3pp, p<0.1) than the matched control group hospitals (Appendix D, Exhibit D-6). This is in contrast to PY1 and PY2, where we did not find evidence that the proportion of patients discharged from the hospital to various PAC settings changed differently between CJR and the matched control hospitals.

The reductions in SNF days, HH visits, and HH PT/occupational therapy (OT) visits achieved by non-opt-in hospitals during PY1 and PY2, when they were in the CJR model, persisted during PY3 and PY4, when they were no longer in the model. For CJR patients with at least

one stay, average SNF days decreased by 2.2 days relative to the matched control group patients during the first two performance years (p<0.01; Appendix D, Exhibit D-6). For patients with at least one HH visit, there was a relative decrease in both HH visits (0.8 visits, p<0.01) and in HH PT/OT visits (0.6 visits, p<0.01). All three of these impacts remained at similar magnitudes during PY3-4.

During the first two performance years, CJR patients at non-opt-in hospitals discharged to a SNF or HHA had less improvement in functional status during their PAC stay than the matched control group. Among patients first discharged to a SNF during the first two performance years, the proportion of CJR patients whose mobility improved during their SNF stay decreased by 4.5pp relative to the control group (p<0.01), and the proportion of CJR patients with improved toilet use decreased by 5.7pp relative to the control patients (p<0.01) (Exhibit 48). During PY3-4, the impact on these measures decreased in magnitude and were no longer statistically significant.

Exhibit 48: Patients discharged from non-opt-in CJR hospitals had less improvement in functional status than control patients

First PAC discharge setting	Measure	PY1-2 DiD	p-value	PY3-4 DiD	p-value
IRF	IRF mobility score	-0.2	p=0.48	0.2	p=0.58
SNF	Improved mobility	-4.5	p<0.01	-2.9	p=0.13
	Improved toilet use	-5.7	p<0.01	-3.1	p=0.15
	Without moderate to severe pain[a]	0.4	p=0.81	1.9	p=0.26
HHA	Improved ambulation/locomotion	-0.8	p=0.17	-1.0	p=0.16
	Improved bed transferring	-1.5	p<0.10	-2.1	p<0.10
	Reduced pain[b]	0.9	p=0.35	0.1	p=0.91

Source: CJR evaluation team analysis of Medicare claims and enrollment data, IRF-PAI data, MDS data, and OASIS data for episodes initiated in 2012 through 2014 that ended between April 2012 and March 2015 (baseline) and episodes initiated during or after April 2016 that ended by September 2019 (intervention).

Notes: The estimates in this exhibit are the result of a DiD model. DiD estimates that are significant at the 1%, 5%, or 10% significance level are indicated by red, orange, or yellow shaded cells, respectively.

DiD = difference-in-differences, HHA = home health agency, IRF = inpatient rehabilitation facility, IRF-PAI = Inpatient Rehabilitation Facility Patient Assessment Instrument, MDS = Minimum Data Set, OASIS = Outcome and Assessment Information Set, PAC = post-acute care, PY = performance year, SNF = skilled nursing facility.

[a] The pain measure for those initially discharged to a SNF was not risk adjusted following the specifications of the MDS 3.0 Quality Measure for short-stay patients used in the CMS Nursing Home Five-Star Rating System.

[b] We cannot be certain that there is no impact of the model, because this outcome failed parallel trends tests (Appendix K). Parallel trends is an assumption that underlies our methodological approach, and without it we do not necessarily believe the control group acts as an accurate representation of what would have occurred in CJR hospitals in the absence of the CJR model. Please see Appendix C (Section III.C.1.c) for additional details.

During the first two performance years, CJR patients at non-opt-in hospitals first discharged to an HHA had a relative decline in bed transferring. The proportion of CJR patients with improved bed transferring decreased by 1.5pp relative to the matched control group (p<0.10). This impact persisted during PY3-4, when the non-opt-in hospitals were no longer participating in the model.

CJR and control patients first discharged to an IRF had similar improvement in functional status during their IRF stay during PY1-2 and PY3-4.

d.   Conclusion

During the first four performance years, opt-in CJR hospitals had a relative reduction in episode payments (3.2%), primarily due to reductions in SNF payments (10.5%). Non-opt-in CJR hospitals, on the other hand, only had a relative reduction in episode payments of 1.7% while participating in the model (PY1-2).[73] These impacts for opt-in and non-opt-in CJR hospitals were both less than those achieved by mandatory CJR hospitals, which by construction of the model are located in MSAs with higher average historical payments, (5.2% relative reduction in episode payments, 13.7% relative reduction in SNF payments, Section II.A.1 and Section II.A.3, respectively).

For non-opt-in hospitals, we did not find evidence of a continued reduction in episode payments after they stopped participating in the CJR model, and there are indications that they may have increased the use of more intensive PAC settings. On the other hand, reductions in the length of SNF stays among patients receiving SNF care persisted after the hospitals stopped participating in the CJR model, which could be an indication of a more lasting impact. While these results provide unique information on the potential for lasting effects of a mandatory APM, it should be noted that generalizability is limited because non-opt-in CJR hospitals are characterized by being located in MSAs with lower average historical episode payments and also by declining to continue their participation in the CJR model.

Lastly, for both opt-in and non-opt-in CJR hospitals, patients discharged to PAC had less improvement in their functional status during their PAC stay than patients at matched control hospitals. This is in contrast to our findings from mandatory CJR hospitals, for which CJR and control patients exhibited similar improvement in functional status (Section II.A.5). We observed no relative changes in claims-based measures of quality, except for one measure indicating a relative improvement for opt-in CJR hospitals. These results could be an indication of hospitals reducing necessary care.

B. Financial Risk or Opportunity

1.   What factors were associated with receiving reconciliation payments under the CJR model?

The underlying assumption of the CJR model is that the opportunity to receive reconciliation payments or to avoid repayments incentivizes CJR participant hospitals to invest in care redesign and coordination to increase the efficiency and quality of care provided to LEJR patients.

---

[73]   While non-opt-in CJR hospitals participated in the model (PY1-2), there was a relative reduction of 2.2 SNF days (8.5%, p<0.01), but the change in SNF payments was not statistically significant (-2.6%, p=0.34) (Appendix D, Exhibits D-5 and D-6).

To understand if certain types of hospitals were more likely to financially gain or lose under the CJR model we investigated the relationship between hospital characteristics and their reconciliation payments. Previously, we found that mandatory hospitals received higher reconciliation payments if they were not-for-profit, had higher LEJR volume, had higher quality of care, and served less complex patients.[74]

In performance year 4, the target price used to determine reconciliation payments is based entirely on regional average historical payments, rather than a blend of the historical regional average and hospital-specific amounts as in earlier performance years. This evolution of the CJR model generally decreased target prices for mandatory CJR hospitals because average historical payments across the larger regions, which were defined as Census divisions, diluted the average historical payments of the mandatory hospitals, which were in the 34 historically highest payment MSAs. As a result, we hypothesized that fewer hospitals would receive reconciliation payments in PY4. In addition, we hypothesized that the level of the reconciliation payments and the repayments would be higher because the stop gain and stop loss limits were raised to 20%.[75]

a. Key Findings	$

- In PY3 and PY4, a lower proportion of CJR participant hospitals received reconciliation payments than in previous years, which is consistent with model design changes that increased the share of the target price that was based on regional average episode payments and reduced the share based on hospital-specific payments.

- The average net reconciliation payment per episode was higher for hospitals that served less complex patients, had higher composite quality scores, and had higher LEJR volume.

---

[74]   Centers for Medicare & Medicaid Services. Comprehensive care for joint replacement model - third annual report. https://innovation.cms.gov/data-and-reports/2020/cjr-thirdannrpt. 2020:107-113

[75]   The stop gain/loss limits were 5%/0% in PY1, 5%/5% in PY2, and 10%/10% in PY3.

b. Methods

For each mandatory CJR hospital with at least 20 episodes, we examined reconciliation payments by performance year. We investigated the relationship between hospital characteristics and the average per episode amount the hospital either received as a reconciliation payment or needed to repay to Medicare in each performance year using a multivariate regression model.[76] Patient characteristics, hospital quality performance in the CJR model, and LEJR volume, averaged at the hospital level, were performance year specific. We categorized continuous variables based on a median split of their values to

Acronyms	
CI	confidence interval
DSH	disproportionate share hospital
FY	fiscal year
HCC	hierarchical condition category
IPPS	inpatient prospective payment system
LEJR	lower extremity joint replacement
MSA	metropolitan statistical area
MS-DRG	Medicare-Severity Diagnosis Related Group
NPRA	net payment reconciliation amount
PAC	post-acute care
POS	provider of services
PY	performance year

compare hospitals in the bottom half of the distribution to hospitals in the top half. Additional information about the methods is available in Appendix C, Section VIII.

c. Results

Half of mandatory hospitals received reconciliation payments in PY3 and PY4 (51% in PY3 and 50% in PY4; Exhibit 49) when the target price was based largely or fully on the regional average. In contrast, 58% and 69% of mandatory hospitals received reconciliation payments in PY1 and PY2, respectively, when the target price was more heavily weighted to hospital-specific historical payments.[77] The target price decreased from PY2 to PY3 for 91% of hospitals. For those hospitals, the average reduction in the target price was 7% or $1,743 for MS-DRG 470 elective episodes. From PY3 to PY4, the target price decreased for 75% of mandatory hospitals, and for those hospitals, the average reduction in the target price was 4% or $999 for MS-DRG 470 elective episodes. As a result, hospitals needed to reduce their episode payments more in later years of the CJR model to come below the target price and receive reconciliation payments.

---

[76] The regression model included PY, hospital characteristics (ownership, bed count, affiliation with medical school, Census region, disproportionate share hospital (DSH) patient percentage, ever BPCI LEJR participation, CJR performance quality category, hospital historical payments in relation to PY target price, and LEJR volume) and episode characteristics (average Hierarchical Condition Category (HCC) score; percent female, age 80 or older, dual eligible, Black or African American, disabled, MS-DRG 470 elective, and with prior institutional stay). We calculated the potential repayment amount for PY1 and used it in our analysis for consistency across performance years.

[77] The regional average comprised one-third of the target price in PY1 and PY2 and two-thirds in PY3. In PY4 and PY5, the target price is based completely on the regional average.

Exhibit 49: Half of mandatory CJR hospitals received reconciliation payments in PY4 when the target price was based on regional average payments
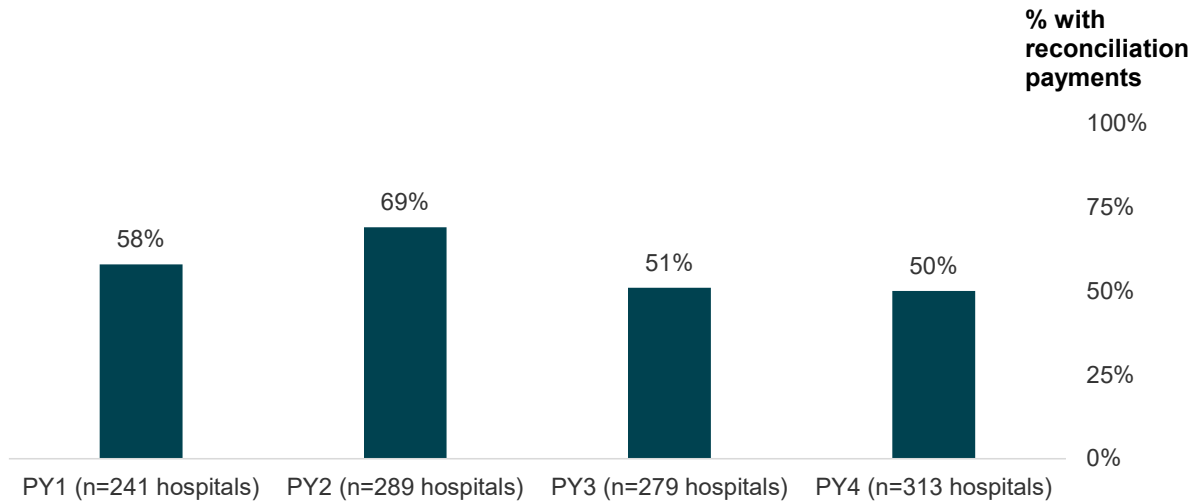
Source: CJR evaluation team analysis of CJR payment contractor NPRA data for CJR participating hospitals in PY1 (episodes starting on or after April 2016 and ending on or before December 2016), PY2 (episodes ending between January and December 2017), PY3 (episodes ending between January and December 2018), and PY4 (episodes ending between January and December 2019).

Notes: Mandatory CJR hospitals with at least 20 episodes in the PY were included, which accounts for the variation in sample size across PYs. Hospitals that did not receive reconciliation payments included those with average episode payments above their quality-adjusted target prices and those that were ineligible due to poor quality.

NPRA = net payment reconciliation amount, PY = performance year.

Factors associated with net reconciliation payments

The average net reconciliation payment per episode varied by performance year, hospital average historical episode payments, hospital average patient complexity, quality performance, LEJR volume, and other hospital characteristics. Appendix L provides detailed result tables from this analysis.

The average net reconciliation payment per episode was lower in PY3 and PY4 than in PY1 ($215 and $349 less, both $p<0.05$), holding all other variables constant (Exhibit 50). In PY3 and PY4, the quality-adjusted target price calculation was weighted more heavily toward the regional average episode payment, which reduced the target price for most hospitals relative to earlier performance years.[77] This is because the larger Census divisions used to calculate the regional average episode payment amounts tend to include more lower cost hospitals than the historically high payment MSAs where the mandatory CJR hospitals were located.

**Exhibit 50: Mandatory CJR hospitals had lower average net reconciliation payments in
PY3 and PY4, compared to PY1**

Difference in average reconciliaton payment per episode, compared to PY1
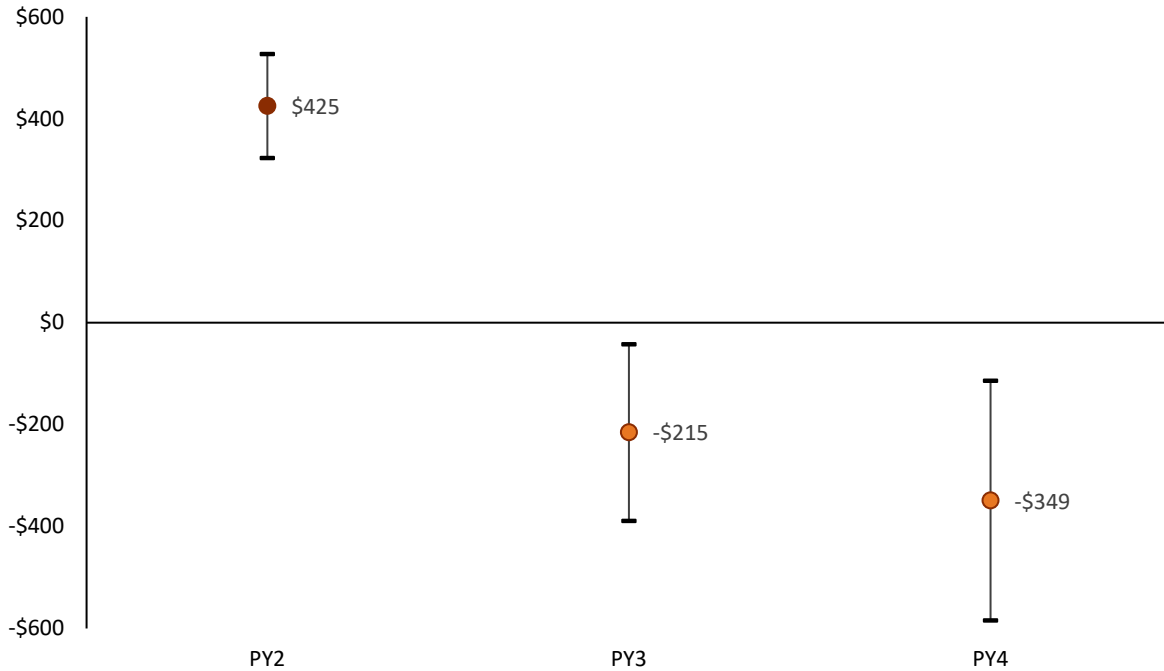
Source: CJR evaluation team analysis of CJR payment contractor NPRA data for mandatory CJR participant hospital(s) in PY1
(episodes initiated during or after April 2016 that ended by December 2016), PY2 (episodes ending in 2017), PY3
(episodes ending in 2018), and PY4 (episodes ending in 2019).

Notes: Multivariate generalized linear regression model was used to identify factors related to average net reconciliation
payment per episode that are significant at the 1%, 5%, or 10% significance level as indicated by red, orange, or yellow
shaded cells, respectively. The whiskers represent 90% confidence intervals.

NPRA = net payment reconciliation amount, PY = performance year.

Approximately one-quarter of hospitals had average historical payments below their quality-
adjusted target price at the start of the performance year and 75% of these hospitals received
reconciliation payments. On average across the four performance years, hospitals with
historical payments below their quality-adjusted target price at the start of the year received
$525 more per episode than hospitals with historical payments at or above their target price
(p<0.10, Exhibit 51). This implies that some mandatory CJR hospitals may not have needed to
reduce their episode payments to receive reconciliation payments because of their historically
low episode payments.

Exhibit 51: Hospitals with average historical payments below their quality-adjusted target price had higher average net reconciliation payments

Measure	Categories	Reference category	Difference in average reconciliation payment per episode from reference group	90% CI	p-value
Hospital historical average payments in relation to the PY target price	Started the PY with hospital historical average payments below the PY target price	Above the target price	$525	[$307 to $742]	<0.01

Source: CJR evaluation team analysis of CJR payment contractor target price data for mandatory CJR participant hospital(s) in PY1 (episodes initiated during or after April 2016 that ended by December 2016), PY2 (episodes ending in 2017), PY3 (episodes ending in 2018), and PY4 (episodes ending in 2019).

Notes: Multivariate generalized linear regression model was used to identify factors related to average net reconciliation payment per episode that are significant at the 1%, 5%, or 10% significance level as indicated by red, orange, or yellow shaded cells, respectively.

CI = confidence interval, PY = performance year.

Hospitals that served less complex LEJR patients had higher average net reconciliation payments per episode (Exhibit 52). Patient complexity was measured across several dimensions, with low and high complexity defined at the median. Hospitals with less complex patient populations, as defined by average hierarchical condition category (HCC) scores, age, proportion of dually eligible patients, and proportion of patients with prior institutional PAC stays, had higher average net reconciliation payments, holding all other variables constant. On average, hospitals with lower average HCC scores (below the median of 1.59) received $526 more per episode than hospitals with higher average HCC scores (p<0.01).[78] Hospitals with lower proportions of patients 80 or older, dually eligible for Medicare and Medicaid, or with institutional PAC stays in the six months before the LEJR had higher average reconciliation payments per episode than hospitals with higher proportions of patients with these characteristics ($200, p<0.10; $320, p<0.05; and $240, p<0.05, respectively). (See Section II.A.6b for further discussion about these patient characteristics and changes in patient mix under the CJR model.)

---

[78]   The HCC score was developed to predict costs in the coming year for a given Medicare patient, compared to the average Medicare patient. A Medicare beneficiary with a HCC score of 1.59 is predicted to have health care costs 59% greater than the average Medicare patient in the coming year.

Exhibit 52: Hospitals that served less complex LEJR patients had higher average net reconciliation payments

Measure	Categories	Reference category	Difference in average reconciliation payment per episode from reference group	90% CI	p-value
HCC score	Below median average HCC score for PY episodes	Above median (>1.60)	$526	[$305 to $747]	<0.01
Age	Below median percent of PY episodes age 80 years or older	Above median (>26.9%)	$200	[$24 to $377]	<0.10
Dual eligibility	Below median percent of PY episodes dual eligible	Above median (>11.2%)	$320	[$105 to $535]	<0.05
Prior institutional stays	Below median percent of PY episodes with prior institutional stays	Above median (>4.9%)	$240	[$55 to $425]	<0.05

Source: CJR evaluation team analysis of Medicare claims and enrollment for mandatory CJR participant hospital(s) in PY1 (episodes initiated during or after April 2016 that ended by December 2016), PY2 (episodes ending in 2017), PY3 (episodes ending in 2018), and PY4 (episodes ending in 2019).

Notes: Multivariate generalized linear regression model was used to identify factors related to average net reconciliation payment per episode that are significant at the 1%, 5%, or 10% significance level as indicated by red, orange, or yellow shaded cells, respectively.

CI = confidence interval, HCC = Hierarchical Condition Category, PY = performance year.

The CJR model rewards hospitals with higher quality through a composite quality score that reduces the effective discount percentage applied to the target price at reconciliation. As intended, hospitals with higher quality scores had higher average net reconciliation payments per episode (Exhibit 53). Compared to hospitals in the "below acceptable" quality category, hospitals with "acceptable," "good," or "excellent" quality received $629, $1,001, and $1,198 higher average reconciliation payments per episode, respectively (all, p<0.01).[79]

---

[79]   Hospitals are not eligible to receive reconciliation payments if they have "below acceptable" quality performance. Hospitals with episode payments below the quality-adjusted target price and "below acceptable" quality had reconciliation values of $0.

Exhibit 53: Average net reconciliation payment per episode was related to quality performance

Measure	Categories	Reference category	Difference in average reconciliation payment per episode from reference group	90% CI	p-value
PY performance quality category	Acceptable	Below acceptable	$629	[$360 to $898]	<0.01
	Good		$1,001	[$801 to $1,201]	<0.01
	Excellent		$1,198	[$931 to $1,465]	<0.01

Source: CJR evaluation team analysis of CJR payment contractor NPRA and quality performance data for mandatory CJR participant hospital(s) in PY1 (episodes initiated during or after April 2016 that ended by December 2016), PY2 (episodes ending in 2017), PY3 (episodes ending in 2018), and PY4 (episodes ending in 2019).

Notes: Multivariate generalized linear regression model was used to identify factors related to average net reconciliation payment per episode that are significant at the 1%, 5%, or 10% significance level as indicated by red, orange, or yellow shaded cells, respectively.

CI = confidence interval, NPRA = net payment reconciliation amount, PY = performance year.

Hospitals with higher LEJR volume had higher average net reconciliation payments (Exhibit 54). Compared to hospitals with less than 15 episodes per quarter, hospitals with 15 to 49 episodes received $560 more per episode and hospitals with 50 or more episodes received $852 more per episode in a given performance year (both, p<0.01).[80]

Exhibit 54: Average reconciliation payment per episode was related to LEJR volume

Measure	Categories	Reference category	Difference in average reconciliation payment per episode from reference group	90% CI	p-value
PY average quarterly volume	15-49 episodes	<15 episodes	$560	[$346 to $774]	<0.01
	50 or more episodes		$852	[$569 to $1,135]	<0.01

Source: CJR evaluation team analysis of CJR payment contractor NPRA data for mandatory CJR participant hospital(s) in PY1 (episodes initiated during or after April 2016 that ended by December 2016), PY2 (episodes ending in 2017), PY3 (episodes ending in 2018), and PY4 (episodes ending in 2019).

Notes: Multivariate generalized linear regression model was used to identify factors related to average net reconciliation payment per episode that are significant at the 1%, 5%, or 10% significance level as indicated by red, orange, or yellow shaded cells, respectively.

CI = confidence interval, LEJR = lower extremity joint replacement, NPRA = net payment reconciliation amount, PY = performance year.

Other hospital characteristics were related to the average net reconciliation payment per episode. Not-for-profit hospitals received an average of $563 more per episode than for profit hospitals (p<0.01, Exhibit 55). Hospitals in the Northeast and South regions had higher average reconciliation payments than hospitals in the West ($459 and $381 more, p<0.05 and p<0.10, respectively). Hospitals with below median bed counts and affiliations with medical schools also received higher reconciliation payments per episode ($265 and $253, both p<0.05,

---

80 Because PY1 was less than a full calendar year of episodes, we used quarterly volume to standardize volume across PYs.

respectively). Hospitals that served a lower percentage of low-income patients, as measured by the disproportionate share hospital (DSH) patient percentage (below the median of 26.9%), received $276 more per episode than hospitals that served a higher percentage of low income patients ($p < 0.05$).

Exhibit 55: Average net reconciliation payment per episode was related to profit status, Census region, bed count, affiliation with a medical school, and DSH patient percentage

Measure	Categories	Reference category	Difference in average reconciliation payment per episode from reference group	90% CI	p-value
Ownership	Not for profit	For profit	$563	[$320 to $807]	<0.01
	Government		$356	[-$28 to $740]	0.13
Census region	Northeast	West	$459	[$97 to $822]	<0.05
	South		$381	[$57 to $706]	<0.10
	Midwest		$222	[-$204 to $649]	0.39
Bed count	Below median bed count	Above median (>258)	$265	[$59 to $471]	<0.05
Affiliation with medical school	Yes	No	$253	[$42 to $464]	<0.05
DSH patient percentage	Below median hospital DSH patient percentage	Above median (>24.2%)	$276	[$60 to $491]	<0.05

Source: CJR evaluation team analysis of December 2016 POS and FY 2016 CMS Annual IPPS data for mandatory CJR participant hospital(s) in PY1 (episodes initiated during or after April 2016 that ended by December 2016), PY2 (episodes ending in 2017), PY3 (episodes ending in 2018), and PY4 (episodes ending in 2019).

Notes: Multivariate generalized linear regression model was used to identify factors related to average net reconciliation payment per episode that are significant at the 1%, 5%, or 10% significance level as indicated by red, orange, or yellow shaded cells, respectively.

CI = confidence interval, DSH = disproportionate share hospital, FY = fiscal year, IPPS = inpatient prospective payment system, POS = Provider of Services, PY = performance year.

d. Conclusion

For the first four performance years of the CJR model, we identified relationships between hospital characteristics and reconciliation payments, while controlling for a variety of market, hospital, and patient characteristics. We found that features of the target price calculation were related to reconciliation payment amounts. We also found that higher volume hospitals had higher average net reconciliation payments. Finally, the relationship between higher quality performance and higher reconciliation payments is consistent with model intent, to reward hospitals that achieve quality metrics.

The target price calculation was designed to encourage historically high-payment hospitals to lower episode payments over time as the target price declines. Starting in PY4, the target price is based completely on the regional historical average. In addition, the 3-year historical period

used in calculating the target price shifts so that by PY4 the period includes performance under the CJR model, which further reduced the target price for many hospitals. In PY3 and PY4, target prices for most hospitals decreased and half of hospitals received reconciliation payments, which was a smaller proportion than observed in the first two years of the model.

Hospitals that started the model with payments below their target prices received higher average net reconciliation payments per episode. These hospitals may not have needed to change their practice patterns to receive reconciliation payments under the CJR model, and if they did make changes they could have been even more financially advantaged under the model.

The relationship between patient complexity and reconciliation payments may indicate that the simple risk stratification methodology based on MS-DRG and fracture status used by CMS to set quality-adjusted target prices did not adequately account for variations in patient complexity that affected episode payments within the four episode categories. This could incentivize hospitals to reduce average patient complexity within each category, making it easier to receive reconciliation payments. As discussed in Section II.A.6b, patient complexity for mandatory CJR hospitals declined relative to the control group during the CJR model. For the three-year extension of the CJR model, target price methodology and other changes will better account for payment differences associated with patient characteristics. Outpatient LEJRs will be included as episodes under the model and site neutral target prices will be adopted. Hospitals will no longer be incentivized to retain the healthier LEJR patients in the inpatient setting to receive reconciliation payments. Further, CMS will incorporate additional adjustments to the target prices to account for the effect of beneficiary age, HCC score, and dual eligibility on payments.

C. CJR Participant Hospital Responses

1. How did the CJR model influence relationships between CJR participant hospitals and orthopedic surgeons?

We surveyed orthopedic surgeons to understand whether and how CJR hospitals influenced their clinical practices. Although hospital participants are financially accountable for LEJR episodes under the CJR model, surgeons may have considerable influence over care redesign. So hospital participants have incentives to engage surgeons in care redesign to improve their performance under the CJR model. Hospitals may develop strategies to ensure surgeons' care redesign and PAC discharge patterns align with the desired response to the CJR model. The survey collected surgeon perspectives about their experiences with the CJR model or mandatory CJR hospitals, and the extent that the model fostered changes in the selection of LEJR patients, clinical care, PAC use, and patient outcomes.

> ### a. Key Findings
>
> - Hospitals provided guidelines or directives to surgeons about patient risk factors to consider when deciding whether to perform an LEJR. Guidelines or directives often related to modifiable health risk factors (e.g., uncontrolled diabetes, obesity, and patient smoking).
>
> - Hospitals provided performance feedback or data to surgeons, which most respondents indicated influenced changes to their care practices.

b. Methods

We identified all surgeons who performed LEJR at mandatory CJR hospitals over a one-year period and applied criteria to identify surgeons who would be knowledgeable about the period between the start of the CJR model in April 2016 and the survey fielding in 2019. We sampled 866 surgeons who met these criteria and emailed or mailed them the survey. We received responses from 249 of 866 sampled surgeons (29%). The survey asked about their relationships with CJR hospitals and whether hospitals influenced LEJR patient selection and pre-surgery, in-hospital, and post-surgery care during the CJR model

Acronyms	
ACO	accountable care organization
BPCI	Bundled Payments for Care Improvement
ED	emergency department
FFS	fee-for-service
ICS	internal cost savings
LEJR	lower extremity joint replacement
LOS	length of stay
PAC	post-acute care
PCP	primary care provider
SNF	skilled nursing facility
TKA	total knee arthroplasty
VBP	value-based payments

implementation period (Appendix M). The survey also included questions about respondents' perspectives and experiences regarding patient outcomes, performance monitoring, gainsharing, and outpatient TKA. In this section, we incorporated previous findings from telephone interviews and site visits or impact estimates when they added relevant context. More detail on the survey is available in Appendix C (Section I.B.2 and Section X).

c. Results

As reported, CJR participant hospital representatives indicated that they collaborated with physicians to reduce spending and improve quality of care.[81,82,83] Hospital interviewees indicated the importance of surgeon engagement in responding to the CJR model. Interviewees described encouraging orthopedic surgeons to discharge patients directly home instead of to a SNF, or

---

[81] Centers for Medicare & Medicaid Services. Comprehensive care for joint replacement model - first annual report. https://innovation.cms.gov/files/reports/cjr-firstannrpt.pdf. 2018.

[82] Centers for Medicare & Medicaid Services. Comprehensive care for joint replacement model - second annual report. https://innovation.cms.gov/files/reports/cjr-secondannrpt.pdf. 2019.

[83] Centers for Medicare & Medicaid Services. Comprehensive care for joint replacement model - third annual report. https://innovation.cms.gov/data-and-reports/2020/cjr-thirdannrpt. 2020.

working with surgeons to standardize implants, order sets, clinical pathways, and pain management practices.

Hospital guidelines or directives for surgeons

Survey respondents indicated that they received guidelines or directives from hospitals about patient risk factors to consider in determining whether to perform an LEJR (Exhibit 56). Over half of respondents reported that hospitals provided guidelines or directives that they should consider uncontrolled diabetes, obesity, or patient smoking; fewer than half indicated that these guidelines or directives changed during the CJR model. Roughly half of respondents considered this guidance at least somewhat important when considering whether to perform an LEJR. Fewer respondents indicated that hospitals provided guidelines or directives about the environmental health risk factors we asked about (e.g., lack of caregiver support or transportation, or safety of the home environment).

Exhibit 56:   Hospital guidelines or directives provided to surgeons more often related to patients' modifiable health risk factors

		Guidelines or directives		
Risk factor type	**Risk factor**	**Provided by hospitals**	**Changed during the CJR model**	**At least somewhat important for decision-making**
Modifiable health risk factors	Uncontrolled diabetes	66%	45%	64%
	Obesity	56%	42%	53%
	Patient smoking	51%	30%	48%
Environmental risk factors	Lack of caregiver support	34%	24%	31%
	Lack of transportation	22%	12%	20%
	Safety of the home environment	38%	21%	36%

Source: CJR evaluation team analysis of the orthopedic surgeon survey data, fielded between August and October 2019.

Note: Estimates are weighted to adjust for survey non-response.

Surgeons reported different approaches to treating patients with risk factors (Exhibit 57). Over half of respondents indicated they postponed surgery or provided the patient with instructions about how to address a modifiable health risk factor, while fewer took similar actions to address an environmental risk factor. Respondents also commonly reported referring patients to a primary care provider or specialist to address a modifiable health risk factor. Respondents reported addressing environmental risk factors by planning to discharge a patient to an institutional PAC setting.

Exhibit 57: Surgeon actions addressing risk factors varied

Risk factor type	Postpone surgery	Provide instructions about how to address the risk factor	Plan to discharge to institutional PAC	Refer to a PCP [a]	Refer to a specialist [a]
Modifiable health risk factors	79%	62%	7%	79%	70%
Environmental risk factors	50%	61%	52%	-	-

Source: CJR evaluation team analysis of the orthopedic surgeon survey data, fielded between August and October 2019.

Notes: Estimates are weighted to adjust for survey non-response.

PAC = post-acute care, PCP = primary care provider.

[a] Respondents were asked to mark all that apply after reviewing a list of possible actions. A "-" indicates that the action was not a response option for environmental risk factors.

A majority of respondents reported changes to hospital LOS (88%), pain management (82%), anesthesia protocols (79%), or simplifying wound dressings (63%) coincident with the CJR model. Over half of surgeons reported that hospital-provided guidelines or directives at least somewhat influenced changes to hospital LOS (68%), pain management (51%), or anesthesia protocols (51%) (Exhibit 58). These guidelines and changes implemented by CJR participant hospitals may have been intended to reduce costs during the inpatient stay through improved care processes and shorter stays. These intentions align with the goals of the CJR model, however prior reports noted that the CJR model was only one of a number of factors influencing interactions between CJR participant hospitals and other providers, including orthopedic surgeons.

Exhibit 58: Changes to care processes occurred during the CJR model, though the influence of hospital guidelines or directives varied

Source: CJR evaluation team analysis of the orthopedic surgeon survey data, fielded between August and October 2019.

Notes: Estimates are weighted to adjust for survey non-response.

LOS = length of stay.

Discharge decision-making

In prior reports, we noted that hospital interviewees said they encouraged surgeons to start discharge planning earlier or support patient attendance at pre-surgical education classes, implemented interdisciplinary rounding, or improved coordination between emergency room physicians and surgeons. Based on this survey, orthopedic surgeons indicated changes in care practices consistent with what the hospitals indicated they were requiring.

Most survey respondents (83%) reported that the proportion of LEJR patients they recommend for discharge to institutional PAC has decreased coincident

> **Surgeon experience with outpatient TKA:**
>
> TKAs were removed from the inpatient only list in 2018 and outpatient procedures are not episodes in the CJR model. Less than 40% of respondents (38%) performed outpatient TKA at the time of the survey. Hospital guidelines or directives influenced surgeon decisions regarding inpatient or outpatient surgical setting for almost half (48%) of respondents performing outpatient TKA, while the other half (52%) indicated guidance was either not influential or not provided.

with the CJR model. Over half of respondents (63%) indicated that hospital guidelines or directives for surgeons regarding discharge destination changed during the CJR model, and guidelines or directives at least somewhat influenced most respondents' (67%) decisions. In addition to hospital guidelines or directives, surgeons also reported that clinical factors (e.g., ability to ambulate or transfer after surgery) and environmental factors (e.g., safety of the

patient's home environment) are important to consider in determining discharge destination, and these factors did not change under the CJR model.

Data sharing and performance feedback

Hospitals have used the CJR data provided by CMS during model participation to share information with hospital leadership, surgeons, or PAC providers. Sharing CJR performance data was noted previously as a critical strategy for engaging physicians in hospital activities related to the CJR model. A prior report also noted the data was particularly effective when working with surgeons to shift discharge destinations to lower intensity settings, and the CJR model was influential in the decision to implement data-sharing with surgeons.

The majority of survey respondents (81%) reported receiving performance feedback from hospitals regarding their LEJR patients. Of those receiving feedback, most noted it pertained to readmissions (93%), institutional PAC use (77%), or patient satisfaction or care experience (76%). Respondents also acknowledged receiving data regarding ED visits (61%), total episode cost of care (51%), or patient-reported outcomes apart from satisfaction or care experience (45%). Of the respondents who received performance feedback from hospitals, most (85%) indicated the metrics influenced them to modify their care practices.

> **Surgeon experience with value-based payments (VBP):**
>
> Prior reports indicated surgeons with past VBP experience were more active in a hospital's response to the CJR model, providing ideas about care redesign and being more willing to change practice patterns. Roughly half of survey respondents (52%) had prior experience with a Medicare Accountable Care Organization (ACO), BPCI or BPCI Advanced, or a commercial payer model. While we do not know if surgeons' prior VBP experience occurred at CJR participant hospitals, it is possible that these surgeons were more familiar with care redesign strategies, receiving performance feedback, and financial incentives.

Gainsharing

Previously, we reported that financial arrangements between hospitals and orthopedic surgeons as well as the availability of surgeons in the market affected the level and type of control CJR participant hospitals exerted throughout the LEJR episode. CJR participant hospitals may enter into agreements to share financial gains from internal cost savings (ICS) or reconciliation payments with surgeons. Sharing financial gains may help hospitals engage or reward surgeons for their role in controlling costs, though gainsharing arrangements may also include accountability for reconciliation amounts owed to Medicare.

As reported previously, hospital gainsharing with surgeons resulted in increased buy-in for care redesign activities, such as lowering utilization of institutional PAC, referring patients to preferred PAC providers, or standardizing implants. LEJR volume and the supply of orthopedic surgeons influenced the use of gainsharing agreements, and hospital interviewees noted that both ICS and reconciliation payments were shared with surgeons, though none reported sharing responsibility for repayments to Medicare. Hospital interviewees described quality and utilization thresholds for surgeons to meet to share in savings, and some agreements required

surgeon participation in activities related to the CJR model or compliance with the hospital's preferred implant list.

Among survey respondents (249 of 866 sampled surgeons, or 29%), over half indicated that they had a financial gainsharing arrangement in place (50%) or were in the process of implementing a gainsharing arrangement (5%) with one or more hospitals where they performed LEJR on Medicare FFS patients. About one-third (29%) of respondents were interested in a gainsharing arrangement, but did not have one in place. A prior report noted that surgeons responded positively to the formal expression of partnership established through the agreements, and hospital interviewees perceived that gainsharing increased physician engagement, willingness to implement new care processes, and collaboration to lower total episode costs.

d.   Conclusion

The relationships between hospitals and surgeons are important to consider in the context of the CJR model; while the CJR model holds hospitals accountable for LEJR episodes, orthopedic surgeons influence episode costs through the decisions they make and services they provide. Relationships between hospitals and surgeons are mediated by various factors, including the provision of hospital guidelines or directives for patient selection and care processes, surgeon decisions regarding patient discharge disposition, hospitals sharing performance feedback with surgeons to modify behavior, and the presence of financial gainsharing agreements.

2.   How did the CJR model influence relationships between CJR participant hospitals and post-acute care providers?

CJR participant hospitals are financially accountable for total episode payments and quality of care for LEJR episodes. The incentives of the model, therefore, encourage hospitals to coordinate care for LEJR patients across the entire episode, including after patients are discharged from the hospital. Consistent with our estimates of lower SNF payments under the CJR model, hospital respondents indicated that reducing SNF care was among the key objectives of their redesign activities and many indicated they leveraged relationships with PAC providers to influence care after patients were discharged from the hospital. We conducted two rounds of telephone interviews with PAC providers who treated patients who had LEJR surgery at mandatory CJR hospitals – one with outpatient PTs and one with SNF administrators – to better understand how the CJR model influenced the care they provided to LEJR patients and their relationships with participating hospitals.

---

a.   Key Findings

- PT interviewees indicated that hospitals did not request changes to care plans for their LEJR patients. In contrast, SNF interviewees said they changed LEJR patient care in response to hospitals' requests. Most often, SNF administrators indicated that hospitals wanted the SNF to provide information about patients during their SNF stay, reduce SNF LOS, and adjust the frequency and timing of physical therapy.

- Most PT and SNF administrator interviewees indicated that their LEJR patients have greater needs now than before the start of the CJR model. Both PTs and SNF administrators indicated this change was across all of their LEJR patients, not just CJR patients, but only SNF interviewees thought the increase in patient complexity was due, at least in part, to the CJR model.

- Both PTs and SNF administrators indicated that their facility's operations and the care they provided to LEJR patients were influenced by a variety of factors including the CJR model.

b.   Methods

We conducted semi-structured telephone interviews with 32 outpatient PTs and 40 SNF administrators about the care LEJR patients received after their hospitalization and whether CJR participant hospitals influenced how they provided care. In this section, we incorporated previous findings when they added relevant context. More detail on these data sources is available in Appendix C (Section I.B.1 and Section X).

Acronyms	
BPCI	Bundled Payments for Care Improvement
EMR	electronic medical record
FFS	fee-for-service
HHA	home health agency
LEJR	lower extremity joint replacement
LOS	length of stay
PAC	post-acute care
PDPM	Patient Driven Payment Model
PPS	prospective payment system
PT	physical therapist
SNF	skilled nursing facility

c.   Results

Patient Care

Background

Previously we reported that participant hospitals responded to the CJR model by beginning care coordination activities prior to admission to the hospital and leveraging relationships with PAC providers to influence care after discharge from the hospital. Some of these activities were underway prior to the CJR model, while the CJR model was a catalyst

> As reported in the second annual report, 62% of CJR hospital survey respondents indicated that they had implemented a preferred provider network for PAC providers and 72% indicated the CJR model influenced their decision to implement or enhance this strategy.

for other activities.[84,85,86] Hospitals extended patient follow-up for a longer period and developed post-acute care protocols and preferred provider networks. Hospital interviewees discussed post-discharge care coordination activities including patient tracking and follow-up, collaboration with PAC providers, and the use of data from outcomes measurement and analytic activities. The processes, timing, and duration of follow-up activities varied across hospitals.

As discussed in previous evaluation reports, many hospital interviewees identified the "highest quality" SNFs and improved the working relationship with those providers. The intent was to minimize SNF LOS and reduce readmissions through higher quality PAC. While many interviewees noted they had preferred provider lists prior to the CJR model, others described updates to their preferred provider selection process under the model. The sophistication of the selection process for preferred providers varied from a "sense that they did well" to selecting those with the shortest LOS, or the use of algorithms and data metrics. Many hospitals reported relying on the CMS Five-Star Quality Rating Systems.

Hospital interviewees described a wide range of requirements for participating preferred PAC providers, including using patient care protocols, following expectations for communication, reporting, and staffing, and having software that allows read-only access to the hospital's electronic medical record (EMR).

Hospitals often reported it was difficult to guide patients to preferred providers because they honored the choice of patients who often made the selection based on convenience (e.g., proximity to their home) or a recommendation of a friend or family member. Hospital staff made efforts to inform patients about their PAC options and emphasized the preferred PAC providers during pre-surgical educational classes or by identifying the preferred providers at the top of PAC provider lists along with their Star Rating.

---

[84] Centers for Medicare & Medicaid Services. Comprehensive care for joint replacement model - first annual report. https://innovation.cms.gov/files/reports/cjr-firstannrpt.pdf. 2018.

[85] Centers for Medicare & Medicaid Services. Comprehensive care for joint replacement model - second annual report. https://innovation.cms.gov/files/reports/cjr-secondannrpt.pdf. 2019.

[86] Centers for Medicare & Medicaid Services. Comprehensive care for joint replacement model - third annual report. https://innovation.cms.gov/data-and-reports/2020/cjr-thirdannrpt. 2020.

Telephone interview findings

Three-fourths of SNF interviewees described making care process changes for LEJR patients in response to hospital requests. These changes included increased communication with the hospital during the SNF stay, reducing SNF LOS, and the frequency and timing of physical therapy during the SNF stay. The majority of SNF interviewees agreed that these changes were requested, at least in part, due to the CJR model.

> *"Obviously one of the biggest changes we've seen is length of stay pressure, to move people along the continuum in a shorter amount of time. To that end, some of our area referral hospitals have even gone so far as to provide us with guidelines for their elective joints."*
>
> – SNF Interviewee

Some SNF interviewees described explicit LOS cutoffs from a hospital, while others described receiving more informal guidance. SNFs identified several changes they implemented to reduce LOS for LEJR patients, including changes in therapy (e.g. increased frequency and intensity, and earlier initiation), advanced planning of admission and discharge, and patient and caregiver education and expectation-setting. Reported

> *"The whole idea was to have a conversation not just for orthopedic patients but for all patients to shorten the length of stay as much as possible. They didn't give us any specific direction."*
>
> – SNF Interviewee

outcomes of the changes varied with about one quarter of interviewees indicating that the changes in LOS resulted in more patients not feeling ready to go home, while a handful of respondents reported generally positive outcomes associated with the changes. One SNF interviewee stated, "Outcomes are good. The only problem is the more complex patients who aren't quite ready to go home and would benefit from staying longer." Most interviewees also indicated that the number of patients who appeal their discharge increased since 2016 however, only a handful of respondents indicated that appeals were more common among the Medicare FFS patient population than among patients with other insurers.

Nearly all of the SNF interviewees reported that their facility was part of at least one hospital's preferred provider network. The majority indicated a hospital invited their facility to join a network at the start of the CJR model and about half of the SNF interviewees attributed involvement in a network to the model. SNF interviewees' descriptions of hospital expectations for SNF participation in the network varied with some hospitals establishing formal, defined requirements for provider participation while others had loosely or ill-defined expectations. The most commonly described expectations were to reduce or maintain low LOS, maintain regular communication with the hospital, reduce hospital

> *"We do a lot of family education here… So families get educated on transferring and getting patients out to doctors' appointments as well as getting them back in their own environment and what kind of assistance they'll have to provide… So a lot of education comes into play with trying to shorten the length of stay."*
>
> – SNF Interviewee

readmissions, and maintain or meet quality ratings (e. g. Star Ratings). Interviewees indicated that referral volume and maintenance of hospital relationships were the main benefits of participation in a provider network. The most frequently identified challenges associated with network participation included the burden of data reporting or documentation and the need to train staff to meet specific requests, especially because of high staff turnover.

In addition to hospitals, SNF interviewees reported increased coordination with other entities, including HHAs, over the past four years that involved more sharing of patient information, earlier discussions regarding discharge planning, and increased communication post-discharge.

Outpatient PT interviewees did not identify any significant changes to care processes or discharge practices as a result of the CJR model. Most PTs described a highly individualized approach to care that was based on patient needs that remained unchanged due to the CJR model. They said they based care plans for LEJR patients on strength and function, patient goals, and post-operative range of motion. Most did not distinguish between Medicare FFS patients and those covered by other insurers when developing care plans. They also did not feel that hospital expectations influenced care processes or PT discharge decisions. As one PT interviewee described, "We are not a protocol-based clinic. There are main phases we are trying to get through and we all have similar ideas around how we are going to get through them but I don't think that has changed. If [patients] have more pain because we are seeing them earlier, that might change the [physical therapist's] focus but the overall philosophy has not changed."

While most PT interviewees were aware of the CJR model, they could not describe model details (e.g. goals of the model, how it affected outpatient facilities, etc.) and none were able to describe the financial components of the model.

Most PT interviewees indicated they had minimal to no communication with hospitals and were not aware of any hospital expectations for care regarding LEJR patients. PTs reported consistent communication with surgeons to report patient progress towards therapy goals, including unexpected patient outcomes or lack of patient progress, as well as questions regarding therapy plans and timing of discharge from outpatient PT. Most PT interviewees indicated there was no difference in the instructions from surgeons for their Medicare FFS patients and patients with other insurance.

Changes in patient mix

Background

Evaluation findings indicate that hospitals in the CJR model send fewer patients to SNFs and more patients directly home with outpatient PT. As we reported in previous evaluation reports, representatives from hospitals participating in the CJR model that we interviewed indicated that the Medicare data they received as participants in the CJR model was useful in determining how to respond to the model. Frequently, they indicated that PAC utilization represented the largest opportunity to reduce episode payments. Hospital representatives reported initiating discharge planning well before the hospital admission when possible to educate patients about their most

appropriate hospital discharge destination and to identify high-risk patients to optimize health outcomes, which could reduce the need for institutional post-acute care.

Telephone interview findings

Both the PT and SNF telephone interviewees indicated that the LEJR population they served had changed since April 2016, the start of the CJR model, though descriptions of how the population changed and perceived influence of the CJR model varied by provider type.

SNF interviewees noted that their LEJR patients have more comorbidities, have less social support, and are discharged from the hospital earlier than LEJR patients prior to the CJR model. As a result, patients arrive at the SNF requiring more services. As one SNF interviewee described, "They definitely have way more comorbidities and are sicker. The patients we used to get 4 to 5 years ago are going home. The patients we are getting now used to stay in the hospital for a week. Now, we are getting them after one day." SNF interviewees indicated that increased patient needs have impacted their operations and care processes, requiring them to hire and train additional staff (e.g. additional nurses or nurses with more advanced degrees), and to change pain management, therapy frequency and timing, post discharge care and coordination, and communication with hospitals. One interviewee described changes in staffing needs, stating, "Being able to care for patients requires a lot more staffing to deal with the increased needs. And, that is a challenge in and of itself, because there is a shortage of nursing staff. We are receiving the same funding but intensity and needs of the patients have increased." For information about the impact of the model on changes in patient mix, see Section II.A.6b and Appendix J.

> *"The social determinants of health are different. Usually it's the elderly person who is living alone, doesn't have that much support around them, so the acute care hospitals need to discharge them somewhere safe. Over time it has become a difficulty for us to get them home safely. They are only in rehab for a short time. It has been beefing up the services post-discharge from us."*
>
> – SNF Interviewee

PT interviewees indicated that they see patients sooner after surgery than they used to, so patients reported higher pain levels and had more acute needs at the start of therapy. One PT stated, "Patients are beginning outpatient physical therapy earlier post-operatively than they had been [previously]. Because of this, patients are now more acute, face more post-operative complications, and experience more post-surgical pain, swelling, and discomfort." Despite increased acuity at the start of therapy, PTs felt that patients were able to meet functional recovery goals by the time of discharge from outpatient physical therapy. PT interviewees did not attribute the change in patient acuity to the CJR model and instead described the trend as starting prior to the CJR model and reflective of broader changes in LEJR care delivery. Similarly, PTs did not believe that changes in patient needs were unique to, or more prevalent among, Medicare FFS beneficiaries than among other patients.

Influencing factors

Background

We have reported that hospitals considered many factors in developing their response to the CJR model, including cost-benefit analyses, market factors, PAC supply, and health system influence. As noted previously, representatives from most hospitals assessed how much their hospital could gain or lose under the CJR model and whether their potential response would be "worth the effort." Hospitals also leveraged LEJR care pathways developed for other episode-based payment approaches, such as Medicare's BPCI initiative or commercial payer bundles. Hospital staff with relevant prior experience indicated greater capacity to identify areas for improvement and implement care redesign changes to succeed under the CJR model. As reported, hospital representatives considered their hospital's market, complete orthopedic service line, resources, and experience in developing their response to the CJR model. It was often not possible, however, to distinguish the influence of the CJR model on changes to care pathways from broader market conditions that affected decisions about the orthopedic service line.

Supply and quality of PAC providers affected how CJR participant hospitals influenced changes to post-acute care pathways. Hospital interviewees identified multiple factors that limited safe patient discharges to PAC providers, including low supply of providers, quality, bed availability, or adequate staffing levels. Many interviewees indicated rural communities in their catchment areas sometimes had only one SNF, and they had little influence over quality of care or discharge from that facility.

Telephone interview findings

Similar to findings from interviews with hospital representatives, PTs and SNF administrators indicated that their operations and LEJR patient care was influenced by a variety of factors in addition to the CJR model.

SNF interviewees said they were under pressure to reduce LOS, not only because of the CJR model, but because of other models, Medicare Advantage, and private payers. Several interviewees also indicated that hospital participation in BPCI or BPCI Advanced influenced changes in communication and coordination with hospitals.

> One SNF interviewee described the intersection of the expectations from the various payer initiatives:
>
> *"The good thing is that they're all looking for the same thing generally. They all want shorter LOS, no returns to the hospital, the patient to be stable enough to move on to the next care setting and be satisfied with their care."*
>
> – SNF Interviewee

About half of the interviewees stated that their facilities implemented changes in response to the new Patient Driven Payment Model (PDPM) under the SNF Prospective Payment System (PPS), including changes in therapy offerings (e.g. shifts to group therapy), changes in

communication with the hospital, and changes to facility operations or management, including documentation and staff training. [87]

Outpatient PTs said that patient progress was the primary consideration in the decision about discharge from PT. Several PTs also indicated that visit limits or the Medicare therapy thresholds influenced decisions regarding the timing of therapy visits and patient discharge.[88]

d. Conclusion

SNF administrators that we interviewed were aware of hospital expectations and consistently reported implementing changes in patient care in response to hospital requests. SNF interviewees attributed some changes to hospital participation in the CJR model, including efforts to reduce length of stay, and noted that patient complexity has increased in part due to the CJR model. Outpatient PTs reported little to no communication with hospitals and that hospitals exerted no influence over the care provided. PTs we interviewed also had little knowledge of the relevance of the CJR model to their practice. Both PTs and SNF administrators indicated that their facility's operations and the care they provided to LEJR patients were influenced by a variety of factors including the CJR model. These findings align with previous evaluation findings that indicated that hospitals focused on efforts to send fewer patients to SNFs and more patients directly home with outpatient PT and leveraged strategies, like preferred provider networks, to influence SNF patient care and length of stay.

---

[87] Effective October 1, 2019, CMS began using the Patient Driven Payment Model (PDPM) case-mix classification system in the SNF Prospective Payment System (PPS) to classify SNF patients in a Medicare Part A-covered stay.

[88] As of 2018, the former Medicare therapy caps now are annual thresholds that physical therapists are permitted to exceed when they append claims with the KX modifier for medically necessary services. This change from the earlier "hard" therapy caps is the result of the Bipartisan Budget Act of 2018 (BBA of 2018), which provides for Medicare payment for outpatient therapy services including PT, speech-language pathology, and occupational therapy services.

III. Discussion and Conclusion

A. Discussion

Over the first four performance years of the CJR model, changes to the fundamental design of the model and broader Medicare policy changes have extended the research questions that can be addressed through its evaluation. This report primarily focuses on the mandatory hospitals in the 34 MSAs with the highest historical payments that were required to participate in the CJR model for this entire period. In addition to findings on the mandatory hospitals, we provide impact estimates for the hospitals in the 33 MSAs that were no longer required to participate starting in PY3, with separate estimates for hospitals that chose to opt-in to the CJR model and those that did not. This report also incorporates our refined methodology to account for the differential response between CJR and control hospitals to the removal of TKAs from Medicare's inpatient-only list.

The CJR model is intended to achieve savings for the Medicare program as participants reduce episode payments more than the reconciliation payments they receive. The calculation of the quality-adjusted target price used in determining reconciliation payments incorporates a 3% to 1.5% discount to help ensure savings, and as the model progresses, the quality-adjusted target price shifts towards regional pricing and includes episode payments achieved under the CJR model in the historical baseline. In the first four performance years for all hospitals that ever participated in the CJR model, estimated Medicare savings was $21 million after accounting for reconciliation payments. Given the wide range around this estimate, however, we cannot conclude that the CJR model resulted in net savings across all hospitals that ever participated. Mandatory CJR hospitals, however, likely achieved savings for Medicare during the first four performance years. In contrast, the opt-in hospitals contributed to significant net losses and it is also likely that the non-opt-in hospitals contributed to net Medicare losses. Medicare savings were lower in PY 3 and 4 because of model and Medicare payment policy changes.

Mandatory CJR hospitals reduced average episode payments in each of the four performance years, while maintaining or improving performance on quality measures. Because the 34 mandatory MSAs had higher average baseline episode payments, the mandatory hospitals may have had greater opportunities to lower episode payments than other hospitals. Their episode payment reductions declined in PY 3 and 4 because mandatory hospitals shifted fewer TKAs to the lower payment outpatient setting. Hospital strategies for reducing LEJR episode payments focused on reducing the use of more intensive institutional PAC services. Even with the decline in PAC use, however, claims- and assessment-based measures indicate that overall quality was maintained or improved for mandatory CJR hospitals. Further, overall patient survey responses indicated similar improvement in functional status and pain from the week before their surgery to after the end of the episode and similar satisfaction with overall recovery, care management, and care transitions. However, CJR respondents required more help from caregivers after returning home, which is consistent with the reduction in PAC use under the CJR model. When we looked specifically at the subset of survey respondents who had LEJR because of a hip fracture, we

found that CJR respondents reported less improvement in functional status than control respondents and also required more help from caregivers at home. Patients with hip fractures may be particularly vulnerable to declines in institutional PAC because their surgeries are not planned, so there is less opportunity for providers to prepare the patients and to coordinate their care. For patients who are Black or African American, eligible for both Medicare and Medicaid (dually eligible), or both Black or African American and dually eligible, there was limited evidence of different impacts of the CJR model on quality of care, functional status, satisfaction, and caregiver help. It would be premature, however, to draw conclusions from this analysis. To fully understand the impact of the CJR model on health equity, it is necessary to understand changes in access to care for subpopulations, in addition to changes in health outcomes. We will continue to evaluate the impact of the CJR model on patients with fracture and other subpopulations that have historically experienced worse outcomes.

This report also presents information about the coordination of care across a range of providers involved in LEJR episodes from a survey of orthopedic surgeons and interviews with SNF administrators and outpatient physical therapists. Orthopedic surgeon survey respondents indicated that some CJR hospitals provided guidelines or directives for patient selection and care processes, some shared performance feedback to modify behavior, and some used financial gainsharing agreements to align incentives. SNF administrators that we interviewed indicated that hospitals set expectations for their CJR patients and requested changes in patient care. Outpatient PTs, however, reported little to no communication with hospitals about expectations for their CJR patients. The consistent message across all of these providers involved in the LEJR episode was that care decisions were influenced by a variety of factors, including the CJR model.

As with all models of this scope and complexity, there could be unintended consequences, which may manifest as hospitals receiving reconciliation payments for reasons other than improving the efficiency or effectiveness of their LEJR episodes. While the CJR model was designed to minimize these opportunities, opportunities may remain. In response to the model, we found no indications that participants were boosting admissions of patients for LEJR or changing their coding practices to lower episode spending. However, for patients with elective MS-DRG 470 episodes, which accounted for about 73% of episodes, the patient population in mandatory CJR hospitals was relatively healthier than in control hospitals during the intervention (PY1-4) than the baseline period. This decline in patient complexity for the mandatory CJR hospitals relative to the control group contributed to the decline in episode payments. As a result, some mandatory CJR hospitals probably received reconciliation payments in part because they treated relatively less complex patients under the model.

As noted, the change in the CJR model beginning in PY3 that allowed hospitals in 33 previously mandatory MSAs to opt-in to the model expands the range of issues that can be explored in the evaluation of the CJR model. The 33 voluntary MSAs had lower baseline episode payments than the mandatory MSAs. Even so, hospitals in the voluntary MSAs reduced episode payments under the model, demonstrating that even in areas with lower payments there were still opportunities to achieve payment reductions. The opt-in hospitals achieved episode payment

reductions in all four performance years. Their reconciliation payments, which were larger than those to mandatory or non-opt-in hospitals, were higher than their payment reductions, so opt-in hospitals contributed to Medicare program losses over the first four performance years. The non-opt-ins achieved payment reductions during the first two performance years when they participated in the model. The decline in episode payments for non-opt-in hospitals, however, did not persist after they stopped participating in the CJR model. Even though their reconciliation payments were lower than those for the mandatory and opt-in hospitals, their reconciliation payments were still higher than their payment reductions. As a result, the opt-in hospitals likely contributed to Medicare program losses over the first four performance years. While these results provide unique information on different outcomes for providers that choose to participate in a model and the potential for lasting effects of a mandatory APM, it should be noted that the opt-in and non-opt-in CJR hospitals are not representative of all potential participants in the CJR model, so these results may not generalize to other APMs.

Over time, target prices under the CJR model were intended to decrease as participants adapted to the incentives of the model. Indeed, in PY3 and PY4, target prices for most hospitals decreased because the regional average episode payments contributed a greater share to the target price. Half of mandatory CJR hospitals received reconciliation payments in PY3 and PY4, which was a smaller proportion than observed in the first two years of the model. At the same time, the evaluation indicates that hospitals with historical payments below their target prices received higher reconciliation payments. These hospitals may not have needed to change their practice patterns to receive reconciliation payments under the CJR model. Hospitals that served less complex patient populations or had higher quality scores also received higher reconciliation payments. The relationship between patient complexity and reconciliation payments may indicate that the simple risk stratification methodology used by CMS to set quality-adjusted target prices based on MS-DRG and fracture status did not adequately account for variations in patient complexity that affect episode payments within the four episode categories. The relationship between higher quality performance and higher reconciliation payments is consistent with model intent, to reward hospitals that focus on quality.

Additional changes in Medicare policy will affect hospitals participating in the CJR model and the impact of the CJR model in the future. Beginning with PY5, Medicare coverage expands to include THAs in the hospital outpatient department and TKAs will be covered in ambulatory surgery centers. Further, TKAs and THAs performed in the hospital outpatient setting will be episodes in the CJR model, under a 3-year extension of the CJR model.[89] In those years, the CJR model's site neutral target prices could affect the share of LEJRs in the hospital outpatient setting, which could affect future impact estimates.

---

[89]   Centers for Medicare & Medicaid Services. Medicare Program: Comprehensive Care for Joint Replacement Model Three-Year Extension and Changes to Episode Definition and Pricing; Medicare and Medicaid Programs; Policies and Regulatory Revisions in Response to the COVID–19 Public Health Emergency; Final Rule 2021:1-81.

B. Considerations

We have employed a robust mixed methods approach that assesses the impact of the CJR model through multiple types of analyses. This approach allows results to be triangulated across data sources and methods, with shortcomings or open questions from one analysis addressed by another. Quantitative results from claims, patient assessments, and patient surveys combined with information gleaned from the orthopedic surgeon survey and provider telephone interviews provide a strong evaluation of the CJR model. Consistency across findings lends strength to our conclusions, while inconsistencies raise questions for further inquiry. The CJR model's mandatory, randomized design mitigates some of the most important concerns that have hampered the evaluation of previous, voluntary episode-based payment models, including selection bias and inability to generalize.[90]

While the results in this report confirm the promise of a mandatory episode-based payment model, several considerations and caveats are important to note. Our evaluation seeks to isolate the impact of the CJR model, however, interactions between the CJR model and other CMS policies and initiatives make it challenging to do so and have required refinements to our methodologic approach. In response to Medicare coverage of outpatient TKAs, CJR participant hospitals shifted a lower share of TKAs to the hospital outpatient setting than control group hospitals, and evidence indicates that the lower share is due to the CJR model. In the third annual evaluation report, we reported two estimates of the impact of the model on average episode payments (one based on inpatient LEJR episodes and the other based on all inpatient and outpatient LEJRs) that were intended to bound the true impact of the CJR model. To improve our impact estimate of the CJR model while taking into account its interaction with this other CMS policy change, in this report, we employed the propensity score weighting method and included *all* control outpatient TKAs in the DiD model, weighted by the hypothetical probability of an outpatient TKA being inpatient if the hospital had been participating in the CJR model.

In October 2018, the BPCI Advanced initiative began. This model also includes LEJR as a clinical episode and aims to reduce payments, while maintaining or improving quality. CJR participant hospitals could not participate in the BPCI Advanced initiative for LEJR clinical episodes; however, hospitals and surgeons in the control group could participate. We found that 46% of mandatory control group episodes that started on or after October 1, 2018 were attributed to the BPCI Advanced initiative. As a result of a high share BPCI Advanced episodes in the control group, the impact of the CJR model on average episode payments and quality of care may be underestimated. To account for this, we included an indicator in our risk adjustment model that flags control episodes attributed to BPCI Advanced.

Of perhaps equal importance, hospitals that participated in the former BPCI initiative contributed nearly three times the number of episodes to the CJR group than to the control group during the performance period, despite contributing nearly the same number of episodes during the baseline.

---

[90] 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.

The differential contribution of episodes by former BPCI hospitals during the performance period is important because, on average, former BPCI hospitals had lower episode payments than non-BPCI hospitals in the performance period, which could contribute to an overestimate of the impact of the CJR model. This finding is in line with interviews from hospitals that stated that former participation in episode-based payment models set them up to swiftly and successfully respond to the CJR model. Further testing to quantify the influence of the differential participation by former BPCI hospitals in the CJR and control groups demonstrated that the differential participation may be overestimating the reductions in average episode payments by roughly $200 per episode. The statistical significance of the impact estimate did not change under these tests, however.

Other factors may lead to an underestimate of the impact of the CJR model. The qualitative analyses indicate that diffusion of CJR care practices to non-CJR hospitals takes place through health systems, which could contaminate our control group. Any influence of the CJR model on control group hospitals in the same health system as CJR participant hospitals would likely result in an underestimate of the CJR model.

The analysis of the telephone interview data describe common themes interviewees discussed when asked about their response to the CJR model and how the CJR model influenced the care LEJR patients received. For interviews, we used purposive sampling to ensure that we captured a range of information on specific topics of interest for each data collection activity. While this sampling approach generates rich data about the population of interest and identifies common themes in providers' responses to the CJR model, it limits generalizability to the broader population.

The response rate of the surgeon survey was 29%. Although responding surgeons were similar to non-respondents on most characteristics, respondents may differ from non-respondents in unobserved ways. Further, this was a cross-sectional survey and did not include a comparison group; thus, the surgeon survey results cannot inform statements about cause and effect.

Our evaluation includes numerous outcomes, which increases the risk that some of our statistically significant findings are due to chance. The strong statistical significance of many of our results and the consistent pattern of results across outcomes implies they are unlikely to be affected by this issue.

In 2020, CMS proposed changes to the CJR model that could affect Medicare savings. Under the current CJR model, it is likely that Medicare savings will continue to decrease into the fifth performance year as more time passes under the outpatient TKA payment policy and as Medicare coverage expands to include THA in the hospital outpatient department and TKA in ambulatory surgery centers beginning in January 2020. In part in response to these changes, CMS published a proposed rule in February 2020 and a final rule in May 2021 that extends the CJR model for an additional three years and expands the CJR episode definition during the

extension to include TKAs and THAs performed in the hospital outpatient department.[91,92] The final rule also makes changes to the target price calculation and reconciliation process that are intended to improve the accuracy of target prices.

In addition, beginning with the fifth performance year, the CJR model evaluation will be affected by the COVID-19 pandemic. There will be fewer CJR episodes during the pandemic as CMS issued guidance stressing the need to avoid elective surgeries.[93] In response to the COVID-19 pandemic, CMS issued interim final rules in April 2020 and November 2020 that extended PY5 of the model through September 30, 2021.[94,95] Further, CMS provides financial safeguards for CJR participant hospitals during the COVID-19 public health emergency. All episodes occurring during the COVID-19 public health emergency will be capped at the target price at reconciliation. After March 31, 2021, CMS will use a more targeted approach and only cap episodes at the target price if they have a COVID-19 diagnosis. The capping of all episodes at the target price during most of PY5 will effectively eliminate repayments for these episodes and the CJR model will likely result in losses during that time.

C. Conclusion

This fourth annual evaluation report demonstrates that the CJR model, which holds hospitals accountable for payments and quality for an episode of care that begins with LEJR surgery, remains a promising approach for reducing episode payments. Through the fourth year of the model, participating hospitals continued to respond to its financial incentives by shifting patients to less intensive PAC settings, resulting in a relative reduction in episode payments. Quality of care was maintained or improved for mandatory hospitals, although CJR patients with a hip fracture reported less improvement in functional status than comparison patients. Additionally, CJR patients, particularly those with a hip fracture, required more caregiver help. For patient subpopulations with historically poorer access to care and health outcomes, there was limited evidence of different impacts of the CJR model on quality of care, functional status, satisfaction,

---

[91] Centers for Medicare & Medicaid Services. Proposed Rule 42 CFR Part 510. Comprehensive Care for Joint Replacement Model, Three-Year Extension and Changes to Episode Definition and Pricing. Federal Register, February 2020. Available at: https://www.federalregister.gov/documents/2020/02/24/2020-03434/medicare-program-comprehensive-care-for-joint-replacement-model-three-year-extension-and-changes-to

[92] Centers for Medicare & Medicaid Services. Medicare Program: Comprehensive Care for Joint Replacement Model Three-Year Extension and Changes to Episode Definition and Pricing; Medicare and Medicaid Programs; Policies and Regulatory Revisions in Response to the COVID–19 Public Health Emergency; Final Rule 2021:1-81.

[93] Centers for Medicare & Medicaid Services. Non-Emergent, Elective Medical Services, and Treatment Recommendations. April 2020. Available at: https://www.cms.gov/files/document/31820-cms-adult-elective-surgery-and-procedures-recommendations.pdf

[94] Centers for Medicare & Medicaid Services. Interim Final Rule. Medicare and Medicaid Programs; Policy and Regulatory Revisions in Response to the COVID-19 Public Health Emergency. Federal Register, April 2020. Available at: https://www.federalregister.gov/documents/2020/04/06/2020-06990/medicare-and-medicaid-programs-policy-and-regulatory-revisions-in-response-to-the-covid-19-public

[95] Centers for Medicare & Medicaid Services. Interim Final Rule. Additional Policy and Regulatory Revisions in Response to the COVID-19 Public Health Emergency. Federal Register, November 2020. Available at: https://www.federalregister.gov/documents/2020/11/06/2020-24332/additional-policy-and-regulatory-revisions-in-response-to-the-covid-19-public-health-emergency

and caregiver help. We will continue to evaluate the impact of the CJR model on patients with fractures and other subpopulations with historically poorer access to care and health outcomes. Nevertheless, for the majority of patients, the CJR model reduced episode payments without compromising quality of care.

The reductions in episode payments likely resulted in Medicare savings for mandatory hospitals. Medicare savings was reduced in the third and fourth performance years, however, because mandatory hospitals shifted fewer TKAs to the hospital outpatient setting than the control group. Furthermore, there are indications that some hospitals may have received reconciliation payments at least in part because they treated a healthier mix of patients. As in prior years, there continues to be evidence that characteristics of the target price calculation may have dampened Medicare savings. While Medicare likely realized savings from mandatory hospitals, Medicare did not realize savings for all hospitals during their participation in the CJR model. Hospitals in the 33 voluntary MSAs that chose to continue participating in the model contributed to Medicare losses, and hospitals that did not choose to continue participating likely also contributed to Medicare losses. Notably, hospitals that stopped participating in the CJR model after the second performance year did not continue to reduce average episode payments. Furthermore, they discharged more patients to a SNF after they stopped participating, although the reduction in the number of days spent in a SNF realized under the CJR model persisted. These novel findings raise questions about the permanence of the effects of mandatory APMs and the entities responsible for change.

These results indicate that a broad range of hospitals can respond quickly to payment incentives, both those provided under the CJR model as well as broader Medicare coverage changes. It will be important to monitor how hospitals respond to Medicare coverage of THAs provided in the hospital outpatient department and TKAs provided in ambulatory surgical centers starting in 2020, which will likely affect the impact of the CJR model. The methodological refinements we used in estimating the impact of the CJR model on mandatory hospitals will serve as the basis for future adjustments to address changes in Medicare coverage.

In future reports, we will deepen our understanding of the impact of the CJR model by refining our estimates of Medicare program savings while adjusting for other policies that affect service use and payments, further investigating the relationship between target prices and savings, and incorporating subgroup analyses. We will continue to monitor for unintended consequences and further explore the variation in patient complexity and its impact on reconciliation payments and Medicare savings. As Medicare coverage of THA expands to include the hospital outpatient department and TKA is covered in ambulatory surgery centers, and as hospitals respond to the COVID-19 pandemic, we will continue to evaluate how broader Medicare policy changes and the evolving health care delivery landscape affect the impact of the CJR model.