LEWINGROUP®
Part of Optum Serve

CMS Comprehensive Care for Joint Replacement Model: Performance Year 6 Evaluation In-Depth Report

Sixth 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**

December 2024

LEWINGROUP®
Part of Optum Serve

CMS Comprehensive Care for Joint Replacement Model: Performance Year 6 Evaluation In-Depth Report

Sixth 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)

This project was funded by the Centers for Medicare & Medicaid Services under contract no. HHSM-500-2014-00033I.

Submitted by:

The Lewin Group, Inc.

Authors:

Lewin: Amanda Markovitz, Carol Simon, Lisa Alecxih, Soumita Lahiri, Ian Morrall, Colin Doyle, Elvedin Bijelic, Niklaus Julius, Maxwell Mindock, Kassandra Fate, Hannah LeManske, Stephen Poskitt, Alexander Johann, Ariana Ackerman, Amanda Heinzerling, Mary-Beth Schaefer, Amy Cowell, James Leckie

Abt Associates: Matthew Trombley, Rayan Joneydi

Acknowledgements:

Lewin: Rachel Henke, Alice Chun, Annie Zhang, and Sean Stuhlsatz

Telligen: Christine LaRocca, Fatima Beltran, Becky Blystone, and Stacy-Ann Whitton

GDIT: Sam Ogunbo, Xiaoyin Zhang, Yan Qian, Colleen Kummet, and Christopher Alleman

December 2024

Table of Contents

Executive Summary

Introduction

Model Goals and Evaluation Approach

Implemented on April 1, 2016, by the Centers for Medicare & Medicaid Services (CMS) Center for Medicare and Medicaid Innovation, the mandatory Comprehensive Care for Joint Replacement (CJR) Model seeks to slow Medicare spending growth by rewarding value rather than volume.

The CJR Model tests whether episode-based payment and quality measurement for lower extremity joint replacements (LEJRs) can lower payments and improve quality. The goal of the CJR Model is for patients to have a safe, effective, and positive recovery experience that is free from complications, while maintaining their freedom of choice in providers and services.

Participating hospitals take on responsibilities for patients receiving an LEJR. These include ensuring that patients receive high-quality, coordinated care by all health care providers from the time of the procedure through recovery, including physical therapy and any other at-home rehabilitation care. Providers work with their patients to develop a plan for recovery, including whether they prefer to recover at home instead of a rehabilitation facility.

The evaluation assessed the impact of the CJR Model (compared with the control group) in Performance Year (PY) 6 on outcomes relevant to model objectives. We used Medicare claims and enrollment data, patient surveys, and case study interviews to evaluate the model's impact on cost, quality, and utilization. The 3-year period (2012 through 2014) before the start of the CJR Model serves as the baseline period for the model. We present the highlights of the evaluation, including a discussion of contextual factors that may have influenced model performance.

Sixth Annual Evaluation Covers the First Year of the Model Extension

In the 2021 Final Rule, CMS implemented multiple changes to the CJR Model:

- Made significant payment design changes, including:
  - Including outpatient episodes
  - Updates to the target price calculation
  - Streamlined reconciliation process
  - Additional flexibilities for gainsharing and downstream distribution payments.
- Returned to full mandatory participation - excluding Section 401 rural status, low volume, and voluntary participant hospitals.
- Extended the performance period by 3 years, from October 2021 to December 2024, to evaluate design updates.

All model design changes took effect in performance year 6. For more details on model changes, refer to the Background section below and the Performance Year 6 Evaluation in-depth Report.

Highlights: Model Impact Findings

Most knee and hip replacements now occur in the outpatient setting. Medicare began covering knee and hip replacements performed in outpatient (OP) setting starting from 2018 (for knee) and 2020 (for hip). CMS included LEJRs performed in OP settings in the model starting in PY 6. Consistent with CMS' goal for site neutrality, inpatient and OP CJR episodes do not have separate prices. The aim was to incorporate OP LEJR procedures in a way that would not incentivize participants to choose a setting based on financial considerations over a patient's level of need. CJR hospitals have shifted toward performing LEJRs in OP settings, although at a slower pace than control hospitals. However, this difference in pace has decreased over time, narrowing the gap between CJR and control hospitals. By the last quarter of PY 6, CJR hospitals performed more than 70% of elective LEJRs in an OP setting. The evaluation included OP episodes starting from PY 6 to ensure the findings are robust and generalizable.

The CJR Model significantly reduced episode payments by shifting patients to less intensive post-acute care settings. During PY 6, compared with the baseline, CJR-participating hospitals had significantly lower average episode payments relative to control hospitals (by $1,012, or 3.5% of baseline). Factors contributing to the lower average episode payments include a reduction in inpatient rehabilitation facility (IRF) payments (of $571), likely due to a large decrease in the proportion of patients first discharged to an IRF (3.9 percentage points [pp]), as well as a large relative increase in the percentage of patients discharged home with home health (3.2 pp). Changes in care patterns for elective LEJRs, representing 88% of episodes, drove the overall findings.

The CJR Model generated net savings of $54.2 million for Medicare in PY 6, marking a return to the pattern of saving from the first four performance years. For the first time in any PY, repayments from hospitals contributed to the savings. This change was likely due to the changes in model target pricing made in PY 6. About half of the hospitals received reconciliation payments, totaling $29.4 million The other half paid back $33.6 million to Medicare. Relative to hospitals that received reconciliation payments, CJR hospitals that made repayments were more likely to be safety-net hospitals (SNHs) and serve a higher proportion of patients from underserved populations.

Hospitals maintained quality of care under the CJR Model. Patients attributed to the CJR and control hospitals experienced similar changes in unplanned readmissions, emergency department visits, mortality, and LEJR complications between baseline and PY 6. Additionally, patients attributed to CJR and control group hospitals who responded to a survey on the quality of their care reported similar improvements in function and mobility, similar levels of satisfaction with their overall recovery, and similar levels of help from their caregivers after returning home. These findings suggest that the CJR Model maintained quality of care while reducing the cost of joint replacement episodes.

Background

The CJR Model is a retrospective bundled payment model that requires hospitals in a set of randomly selected metropolitan statistical areas (MSAs) to participate. The model holds participant hospitals financially accountable for the cost and quality of health care services during and after an LEJR. The payment incentives encourage participant hospitals to coordinate care with the physicians, post-acute care (PAC) providers, and other providers and clinicians involved in the 90-day episodes of care defined by the model rules. The model tests whether episode-based payment and quality measurement can reduce costs, improve quality of care, and advance care coordination.

Acronyms

ACO	Accountable Care Organizations
BPCI, BPCI-A	Bundled Payments for Care Improvement (BPCI) initiative and BPCI Advanced
FFS	Fee-for-Service
IRF	Inpatient Rehabilitation Facility
IPPS	Inpatient Prospective Payment System
LEJR	Lower Extremity Joint Replacements
MSA	Metropolitan Statistical Area
MS-DRG	Medicare Severity-Diagnosis Related Groups
OP	Outpatient
PAC	Post-acute care
PY	Performance Year
THA	Total hip arthroplasty
TKA	Total knee arthroplasty

Model Design

The CJR Model began on April 1, 2016. In 2021, CMS implemented key design changes and extended the model to run through December 31, 2024, representing 8 PYs.[1] The model requires all hospitals in a CJR MSA to participate. The model extension applies to CJR participants in mandatory markets, excluding Section 401 rural status, low volume, and voluntary participant hospitals. The CJR Model's mandatory and randomized design includes a spectrum of hospitals with varying levels of infrastructure, care redesign experience, episode costs, utilization, and market positions, which allows for a broad test of the CJR Model.

LEJR surgeries are primarily for hip replacements (total hip arthroplasty) and knee replacements (total knee arthroplasty) and can be elective or due to fractures.[2] For PYs 1–5, an LEJR episode of care began with the hospitalization of an eligible Medicare fee-for-service (FFS) patient at a hospital paid under the Inpatient Prospective Payment System.[3] Starting in PY 6, the definition of an episode of care includes LEJRs performed in the OP setting. Hospitals are accountable for the cost and quality of the surgery and other health care services during the 90 days after hospital discharge.

CMS provides each participant hospital with preliminary target prices prior to each PY. If episode payments fall below the target price, the hospital can earn a reconciliation payment. If episode payments exceed the target price, the hospital repays Medicare. The target prices represent the

[1] 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, 86 FR 23496 (May 3, 2021) (codified at 42 CFR 510).

[2] Identified using International Classification of Diseases codes listed in https://innovation.cms.gov/Files/worksheets/cjr-icd10hipfracturecodes.xlsx.

[3] Beneficiary was enrolled in Medicare Parts A and B, Medicare was the primary payer (not enrolled in any managed care plan), and beneficiary was not eligible for Medicare based on end-stage renal disease.

average spending within that hospital's region for each of the four Medicare Severity-Diagnosis Related Groups (MS-DRGs) related to hip and knee replacement in the CJR Model.[4]

For the model extension that began in PY 6, CMS made multiple changes to the model design, and how the target prices were calculated. CMS anticipated that the changes would reduce Medicare program expenditures and result in savings over the additional 3 model years while preserving or enhancing the quality of care.[1] CMS aimed to make the target prices more accurate and adaptable by aligning with practice patterns and payment methodology, including OP episodes, and recognizing quality care.

Evaluating the CJR Model in the Context of the Current Health Care Landscape

The PY 6 evaluation's approach to measuring the impact of the model and interpreting the findings accounts for the potential influence of other programs and policies, including their influence on CJR hospitals' approach to care transformation. CMS has enacted policies and launched models that may affect CJR and control hospitals and their markets differently. These include the Bundled Payments for Care Improvement (BPCI) initiative and BPCI Advanced; Accountable Care Organizations (ACOs), such as the Medicare Shared Savings Program; Medicare Advantage (MA) programs; and initiatives to address health equity. In particular, the PY 6 evaluation explores the interaction between participation in ACOs and the CJR Model, including the impact of ACOs on CJR hospitals' approach to care transformation. The evaluation also analyses health equity implications of the model, considering the changing landscape and focus on health equity since the model began.

Hospitals' ability to transform the delivery of care under the model varies based on factors such as hospital characteristics, market and population features, alignment with other value-based care (VBC) initiatives, and relationships with health systems, orthopedic surgeons, and associated providers. The alignment of CJR with other market pressures and VBC initiatives influenced how hospitals responded to the CJR Model. Some hospitals leveraged existing partnerships and standardized care protocols to react quickly to the CJR Model, and others enhanced or streamlined VBC initiatives across the hospital or system.

Changes to Target Price Calculation Starting from PY 6

- **Prices are risk-adjusted** based on age, dual-eligibility status, and count of Hierarchical Condition Categories. These adjustments are **in addition** to the existing 3% discount and adjustments for composite quality score.
- **CMS calculates target prices using the most recent year,** instead of 3 years, of claims data.
- CMS replaced a national adjustment factor with a **retrospective market trend factor** that is applied at the reconciliation stage.
- CMS **updated the quality discount factors** for hospitals with a quality rating of "excellent" or "good," altered the method to calculate spending caps, and removed the use of anchor factor and regional- and hospital-specific anchor weights.

[4] For PY 6, the MS-DRG groups are MS-DRG 469: Major Hip and Knee Joint Replacement or Reattachment of Lower Extremity with Major Complications or Comorbidities (MCC); MS-DRG 470: Major Hip and Knee Joint Replacement or Reattachment of Lower Extremity without MCC; MS-DRG 521: Hip Replacement with Principal Diagnosis of Hip Fracture with MCC; MS-DRG 522: Hip Replacement with Principal Diagnosis of Hip Fracture without MCC.

Past evaluation reports describe how CJR hospitals used a range of enhanced or new initiatives to transform care across the presurgical, hospitalization, and post-discharge care pathways for LEJR patients. Hospitals increased their focus on patient education and patient optimization, where providers identify high-risk patients and mitigate medical and social risk factors for improved outcomes and recovery. Hospitals also provided physical therapy earlier and more often, used data to inform clinical decision-making, and worked with surgeons and PAC provider partners to adopt more efficient practices. These efforts can help shift care away from more expensive settings, such as PAC facilities like skilled nursing facilities (SNFs) and IRFs, and toward less expensive options, such as home health care. Coordination with PAC partners can also reduce patients' length of stay in PAC facilities and limit unnecessary readmissions.

Under the current evaluation, we interviewed hospitals that participated in both a Medicare ACO and the CJR Model. We learned that hospitals used common strategies to respond to the programs, including using data to inform care pathways, strengthening care coordination, and monitoring patient outcomes. Interviewees discussed how participation in an ACO and CJR resulted in more awareness and greater alignment toward VBC among hospital staff. Many hospital- and system-level efforts aimed to align management of VBC initiatives across programs. For example, some hospitals and hospital systems created one VBC management team to help coordinate care across programs, used the same PAC preferred provider network for both programs, and developed a new electronic management system to see patient notes and vital signs and track readmission risk across multiple programs. The CJR hospitals that already participated in an ACO used their ACO experience, data, and resources to guide changes to the LEJR pathway in response to the CJR Model. Other interviewees said that the CJR Model motivated the decision to join a Medicare ACO. These participants noted that their hospital system could leverage the technologies and funding from CJR to implement the ACO.

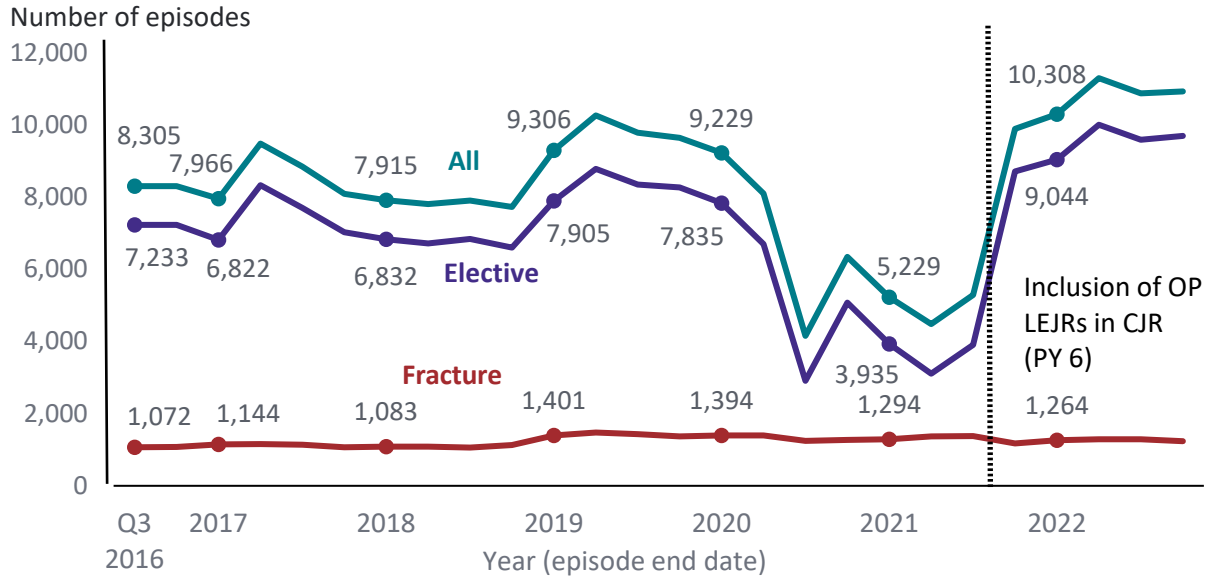
This evaluation also monitors impacts of the model on health equity, and experiences of hospitals transforming care for underserved populations. CMS' Strategic Plan[5] identifies one of its six strategic pillars as advancing health equity by addressing disparities. As part of that strategy, CMS priorities include explicitly measuring the impact of policies on health equity to inform the development of sustainable solutions that close gaps in health and in health care access, quality, and outcomes. Although the original CJR Model design did not provide resources for, incentivize, or require participants to address health equity, changes in hospital practices implemented in response to the model could potentially have differing effects on underserved populations of patients. In addition, initiatives implemented to address CMS' strategic plan can affect care strategies and hospital performances.

[5] Centers for Medicare & Medicaid Services. (2024). *CMS Strategic Plan*. https://www.cms.gov/about-cms/what-we-do/cms-strategic-plan

Summary of Performance Year 6 Report Findings

The CJR Model Population

Exhibit 1: The Number of Episodes Included in the Model at Mandatory CJR Hospitals Has Grown Over Time and Increased Sharply in PY 6 With the Addition of OP LEJRs



Source: CJR evaluation team analysis of Medicare claims and enrollment data for LEJR episodes ending between July 2016 (PY 1) and December 2022 (PY 6).

Notes: CJR = Comprehensive Care for Joint Replacement; LEJR = lower extremity joint replacement; OP = outpatient; PY = performance year.

Over the five quarters of PY 6 (October 2021–December 2022), 320 CJR hospitals performed about 53,000 LEJR surgeries that triggered an LEJR episode of care under the CJR Model.[6]

The CJR Model had more LEJR episodes in PY 6 than in prior years (**Exhibit 1**). About 88% of episodes were for elective procedures. CMS included OP LEJRs as part of the CJR Model for the first time in PY 6, and by the last quarter, more than 70% of the elective episodes for both CJR Model and control hospitals occurred in the OP setting. This increase in LEJRs in an OP setting occurred for both total knee arthroplasty and total hip arthroplasty procedures.

CJR and control hospitals had similar patient populations in both PY 6 and in the baseline period, but between baseline and PY 6, some of the characteristics of patients receiving LEJRs changed in both groups. In PY 6, most patients receiving LEJRs in both the CJR and control group were female (63%) and Non-Hispanic White (85%), and about three-quarters were diagnosed with hypertension and a quarter with diabetes. Between the baseline period and PY 6, the share of LEJR patients dually eligible for Medicare and Medicaid declined by almost one half, the prevalence of obesity doubled, and fewer patients receiving an LEJR had an acute care stay in the 6 months prior or received care in a PAC setting.

[6] The number of hospitals only includes hospitals with CJR evaluation-related LEJR episodes.

Impact of the CJR Model

Payment and Utilization

The CJR Model continues to reduce episode payments, mainly through less use of institutional post-acute care.

CJR hospitals reduced average episode payments. CJR hospitals reduced episode payments by $1,012, or 3.5% of the CJR Model baseline, relative to the control hospitals in PY 6. The reductions were mostly driven by declines in IRF payments of $571, or 25.9% of the CJR Model baseline. We found a similar impact for elective LEJRs, which represent most LEJRs.

For elective episodes, the CJR Model led to a relative reduction of $1,171, or 4.5% of the CJR Model baseline, in average episode payments between the baseline and PY 6, driven by a $410, or 25.1% of baseline, reduction in IRF payments. The model did not have any significant impact on average episode payments for fracture patients. However, CJR Model hospitals had large relative increases in SNF payments and large relative decreases in IRF payments.[7] These two effects appear to offset each other resulting in no relative change in average episode payments.

CJR hospitals sent fewer patients to more intensive PAC settings. CJR hospitals had a relative decline in the proportion of patients first discharged to IRFs for the all-LEJR population (3.9 pp reduction, or 28.3% of the CJR Model baseline) and the elective LEJR population (3.0 pp reduction, or 26.2% of the CJR Model baseline). Among the fracture population, CJR hospitals had a relative increase in the proportion of patients first discharged home with home health under the CJR Model between the baseline and PY 6 (3.2 pp, or a 60.9% increase).

Quality of Care

Our evaluation of claims-based measures and patient-reported outcomes suggests that CJR hospitals maintained quality of care between baseline and PY 6.

CJR and control hospitals experienced similar changes between baseline and PY 6 in claims-based quality of care measures, The rates for all claims-based measures studied (unplanned readmissions, emergency department visits, mortality, and LEJR complications) decreased from baseline to PY 6 – however that decrease was similar for CJR and control hospitals.

Interpreting Impact: Claims-Based Outcomes

- We calculated the **impact of the CJR Model** on payments and quality using a difference-in-differences (DiD) methodology, which subtracts the difference from baseline to intervention for the CJR Model population from the difference for the control population.
- The **percent change from the CJR baseline** is calculated by dividing the DiD estimate by the CJR Model baseline average. This value represents the percent change from the CJR Model baseline that is due to the CJR Model.
- **Average Per-Episode Payments** is the average sum of Medicare fee-for-service payments for all services and items included in the episode. We define payments as standardized allowed amounts, which include beneficiary cost sharing and do not include wage adjustments and other Medicare payment adjustments.

[7] Refer to Annual Report 6 In-Depth Report, Chapter III for limitations on this finding and additional details.

Survey respondents who had elective LEJRs at CJR and control hospitals reported similar experiences in recovery. Patients with LEJRs at CJR and control hospitals reported similar improvement in functional status, indicated similar levels of satisfaction with their overall recovery, and required similar levels of help from their caregivers after returning home. With the inclusion of LEJRs in the OP setting, we also examined differences in outcomes between patients who had an elective LEJR in the inpatient setting relative to those in the OP setting. We found no systemic differences in patient-reported outcomes.

Methods to Examine the Effect of the CJR Model on Function and Care Experience

- We collected responses from a sample of LEJR patients on their status after surgery and recalled status before surgery.

  (The median time patients returned surveys was 37 days after the end of their 90-day post-discharge period.)
- We estimated the effect of CJR as the average difference in responses between CJR and control respondents, adjusting for various characteristics.

Due to the urgent and unexpected nature of hip fractures, we expect these patients to identify having worse functional status after their episode relative to before the episode. Among patients with hip fracture episodes, respondents who had an LEJR at a CJR hospital reported significantly lower levels of decline in using stairs and dependence on a mobility aid from before surgery to after surgery than patients with LEJR at control hospitals. CJR Model respondents with hip fracture also reported higher satisfaction with recovery and care management than did control respondents and were more likely to agree that they were discharged from the hospital at the right time.

Medicare Program Savings and Net Reconciliation Payments

The CJR Model returned to its prior pattern of generating Medicare program savings in PY 6. The savings were large enough to offset losses due to the large payouts to hospitals during the public health emergency. For the first time in any PY, repayments from hospitals contributed to the savings. This change was likely due to the changes in model target pricing implemented in PY 6.

In PY 6, the CJR Model generated statistically significant estimated savings of $54.2 million. The savings may have ranged from $2.0 million to $106.4 million based on 90% confidence intervals. On a per-episode basis, the model saved an estimated $1,017 (with a range from $38 to $1,996). This finding appears to align with CMS' anticipated reduction in expenditures and savings to the Medicare trust fund over the 3-year extension.[8]

Interpreting Medicare Program Savings

Net payment reconciliation amounts (NPRA) are the incentive payments made to CJR hospitals by Medicare, or the net of repayments that CJR participant hospitals make to Medicare for exceeding episode target prices.

Medicare Program Savings (MPS) was calculated as the difference between the estimated change in Medicare spending and net reconciliation payments that CMS made to CJR participants in PY 6:

MPS = Change in episode spending – NPRA

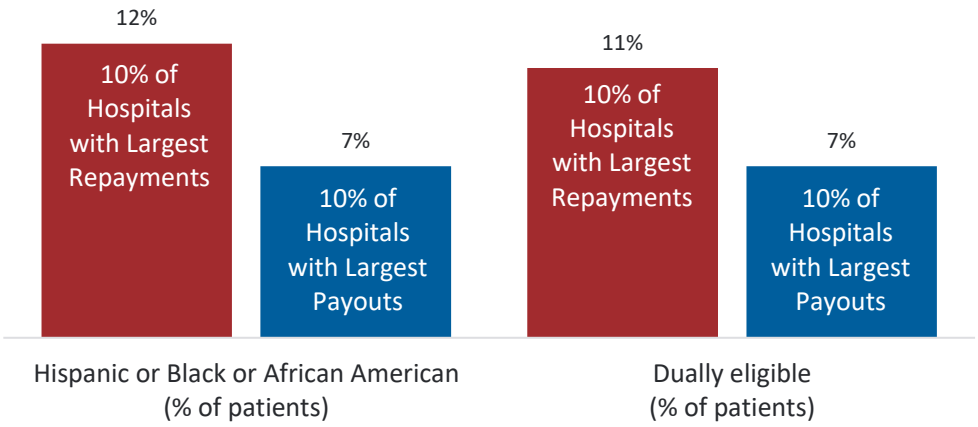
[8] Medicare Program: Comprehensive Care for Joint Replacement Model Three-Year Extension and Changes to Episode Definition and Pricing; Medicare and Medicaid Programs; Policies and Regulatory Revisions in Response to the COVID-19 Public Health Emergency, 86 FR 23496 (May 3, 2021) (codified at 42 CFR 510).

PY 6 was the first PY with an average net repayment to CMS per episode. The average net reconciliation payment per episode was –$78, indicating more net repayments from hospitals to CMS. As a result, the net repayments help contribute to overall Medicare program savings.

In PY 6, about 46% (146) of the hospitals received reconciliation payments from CMS, while 50% of hospitals (161) made repayments to CMS. Thirteen hospitals had no repayment obligation and received no reconciliation payments. The updates to target pricing in PY 6 may have affected changes in the distribution of payments to hospitals compared with previous PYs. Compared with prior years, the hospitals that received the highest relative reconciliation payments received lower reconciliation payments in absolute terms. In contrast, hospitals that made the most relative reconciliation payments were responsible for substantially larger repayments in absolute terms. The 10% of hospitals with the largest repayments collectively repaid $19.3 million. The 10% of hospitals with the largest reconciliation payments collectively received $18.6 million.

The payouts and repayments were not equally distributed across hospitals. Relative to hospitals that received reconciliation, those that made repayments had a higher proportion of underserved patients and a larger proportion were identified as SNH (**Exhibit 2**). Of the 32 (10%) hospitals with the largest repayments, three were identified as SNHs, 11.5% of their patients were either Hispanic or Black or African American, and 10.9% of their patients were dually eligible for Medicaid and Medicare. None of the hospitals receiving reconciliation payments were identified as a SNH.

Exhibit 2: Hospitals That Made Repayments Had a Higher Proportion of Underserved Patients Than Hospitals That Received Payouts



Source: CJR evaluation team analysis of Medicare claims and enrollment data for LEJR episodes ending between July 2016 (PY 1) and December 2022 (PY 6).

We also conducted exploratory analyses to examine if the changes in target pricing affected whether a hospital earned a reconciliation payment or repaid CMS and if hospitals moved in NPRA distribution[9] relative to prior years. Findings indicate relative shifts in NPRA in PY 6:

- **Previously unsuccessful hospitals had limited mobility.** For the 30% of hospitals with the lowest net reconciliation (had either some repayment or small NPRA) in PY 1–5, 44% remained in the bottom 30% of the PY 6 NPRA distribution. A small proportion (11%) of hospitals shifted to the top 30% (received some of the largest reconciliations) in PY 6.
- **Moderately successful and unsuccessful hospitals had more "upward mobility."** For the 40% of hospitals in the middle of the PY 1–5 distribution, many moved to higher-NPRA deciles in the PY 6.
- **A third of previously successful hospitals had notable downward mobility.** One-third of hospitals in the top 30% of the PY 1–5 NPRA distribution (that is, the 30% of hospitals that received the most reconciliation dollars) shifted to the bottom 30% of the PY 6 NPRA distribution (hospitals with the largest repayments).

Health Equity Analyses

CMS' Strategic Plan[10] identifies one of its six strategic pillars as advancing health equity by addressing disparities. As part of that strategy, CMS' 2022–2032 Framework for Health Equity lays out detailed priorities and goals. [11] One of the priorities includes explicitly measuring the impact of policies on health equity to inform the development of sustainable solutions that close gaps in health and in health care access, quality, and outcomes. As part of the approach to advance health equity, CMS also aims to evaluate policies to support safety-net providers, including acute care hospitals. [12]

Health Equity Analyses for PY 6

- **Examined impact of the model for underserved population on:**
  - LEJR volume
  - Cost, utilization, and quality.
- **We conducted impact analyses for four underserved populations:**
  - Black or African American patients,
  - Patients who were dually eligible for Medicare and Medicaid,
  - Patients who are both Black or African American and dually eligible, and
  - Hispanic patients.
- **Conducted interviews with 6 safety-net CJR participating hospitals.**

[9] NPRA distribution refers to deciles based on total NPRA hospitals received (or paid out) ranging from top 10% of hospitals that received the largest reconciliation to bottom 10% of hospitals that had the largest pay out. Shifts in NPRA distribution indicates relative change in hospital NPRA relative to other hospitals and may indicate differential affect of target prices between the hospitals.

[10] Centers for Medicare & Medicaid Services. (2024). *CMS Strategic Plan*. https://www.cms.gov/about-cms/what-we-do/cms-strategic-plan

[11] Centers for Medicare & Medicaid Services. (2023). *CMS Strategic Plan: health equity.* https://www.cms.gov/files/document/health-equity-fact-sheet.pdf

[12] Centers for Medicare & Medicaid Services. (n.d.). *Health equity*. https://www.cms.gov/priorities/innovation/key-concepts/health-equity

Although the CJR Model did not incentivize or require participants to address health equity, changes in hospital practices implemented in response to the model or changes in model design such as including patient characteristics and hospital quality rating in episode target price calculations could have differing effects on underserved populations. For the PY 6 evaluation, we investigated the impact of the model on underserved populations and gathered preliminary perspectives on the experience of safety-net providers participating in the model. Hip fractures, constituting approximately 12% of all LEJRs, present hospitals with limited opportunities for presurgery patient optimization or discretion in scheduling of the surgery. Additionally, fracture patients have different outcomes and model impacts. To avoid conflating the effects on fracture and elective populations, for the health equity analyses, we considered only elective LEJRs.

Model Effects for Underserved Populations

The evaluation examined impact on Hispanic population for first time for PY 6. Findings indicate LEJR volume increased for Hispanic population between baseline and PY 6 in CJR MSAs relative to control. Volume of LEJRs also increased for dually eligible patients. CJR Model reduced costs for underserved populations more than the corresponding contrasting population.

The CJR Model increased LEJR volume for Hispanic and dually eligible patients. In both CJR and control MSAs, underserved populations had substantially lower LEJR volume than their counterparts during both the baseline period and PY 6. The model increased LEJR volume for patients who are Hispanic (11.1% relative to the baseline, $p < 0.01$) and for patients who were dually eligible (7.6% relative to the baseline, $p = 0.11$). There was no difference in LEJR volume for patients who are Black or African American or patients who are both Black or African American and dually eligible between baseline and PY 6 relative to control MSAs.

The CJR Model reduced costs more for underserved populations but had mixed effects on quality. The CJR Model reduced average payments for all patients, but the reduction was substantially larger for patients from underserved populations. Changes in PAC discharge destination were the primary drivers behind these reductions in payments, with a relatively lower proportion of patients discharged to IRFs and SNFs and a higher proportion discharged to home health. Some evidence suggests that the CJR Model led to higher rates of discharge to home without home health for underserved populations.

The CJR Model also led to statistically significant increases in emergency department use for patients from underserved populations, as well as a statistically significant decrease in all-cause mortality for patients who are Black or African American.

Interpreting Impact of Model for Underserved Populations

We estimated the impact of the model on underserved and reference populations using a difference-in-difference-in-differences approach. With this approach, we estimated two effects:

- The impact on the model for each subpopulation as relative change from baseline to PY 6 between CJR and control.
- The difference between the impact for underserved and reference populations, which we refer to as the 'differential impact' of the model.

Underserved populations reported worse functional status and care[13] following LEJRs at both CJR and control hospitals. In both the CJR and control populations, underserved populations had worse experiences relative to non-underserved populations across all measures. Dually eligible patients who had an LEJR at CJR hospitals had worse functional status outcomes and higher likelihood of needing caregiver help (relative to reference patient population) than those who had an LEJR at control hospitals. Black or African American patients who had an LEJR at a CJR hospital had higher levels of satisfaction with care management (relative to reference patients) than Black or African American patients who had an LEJR at control hospitals for all five satisfaction questions.

Safety-Net Hospitals—A Preliminary Perspective on CJR Model Experience

SNHs provide care regardless of patients' insurance status and, thus, typically serve a higher proportion of underserved populations. Empirical studies most often identify SNHs using metrics like disproportionate share percentage, measures of Medicaid caseload, or the amount of uncompensated care that qualifies a hospital for additional Medicare payments.[14,15]

We interviewed representatives of six SNHs to understand how these hospitals responded to the CJR Model and how the care pathways may differ for their patients with unmet nonmedical needs. All six hospitals had a low volume of CJR episodes (21–57 episodes ending in or between Quarter 4 [Q4] 2021 and Q3 2022) and made repayments to CMS under CJR ranging from –$228K to –$921K. Hospitals reported that they serve complex patient populations with high unmet medical and nonmedical needs and that their care transformation efforts focused on addressing social determinants of health (SDOH) at discharge.

Interviewees reported that care transformation strategies that can improve financial success are often not feasible for SNHs. While many CJR hospitals can shift procedures to the OP setting to deliver care at a lower price, as well as reduce institutional PAC use, SNHs said they could not make these changes because their patients' needs often require inpatient procedures and more intensive post-discharge care.

Rather than targeting PAC discharge setting, these hospitals focused on identifying and addressing SDOH. In caring for LEJR patients with unmet nonmedical needs, all six hospitals screened for the social needs for all patients. Half of the hospitals continued these screenings in their discharge planning. Lack of in-home family support, transportation, access to quality PAC services, access to regular primary care, and food security were common unmet nonmedical needs of LEJR patients at SNHs.

SNHs were not optimistic about their financial prospects in the model due to low CJR episode volume and target prices that were "too low" or did not reflect the cost of care for their complex patient population. To better provide care for LEJR patients with unmet nonmedical needs, these

---

[13] Measured effects of CJR are not statistically significant, and the small respondent sample limits our ability to draw conclusive inferences. Refer to In-Depth report for limitations and additional details on findings.

[14] Centers for Medicare & Medicaid Services. (2022). Person-centered innovation – An update on the implementation of the CMS Innovation Center's strategy – supplemental document. https://www.cms.gov/priorities/innovation/data-and-reports/2022/cmmi-strategy-refresh-imp-tech-report

[15] Hefner, J. L., Hogan, T. H., Opoku-Agyeman, W., & Menachemi, N. (2021). Defining safety net hospitals in the health services research literature: a systematic review and critical appraisal. *BMC Health Services Research, 21*(1), 278.

SNHs suggested more funding for community services to address social barriers such as housing, transportation, mental and behavioral health services, and increased coverage for additional in-home care. They also requested changes such as target pricing that reflects the needs of their complex patient populations and higher reimbursement for LEJR procedures to improve their financial performance under the CJR Model.

Potential Unintended Consequences of the CJR Model

The evaluation examined whether the CJR Model had any unintended consequences.

No evidence of changes in the composition of elective LEJR patient population between PY 5 and PY 6. Patient complexity for both CJR and control hospitals increased modestly between the baseline period and PY 6. However, CJR hospitals had a significantly smaller increase, suggesting that on the margin the model may favor some lower-risk patients. In addition, we found no notable differences in changes in the CJR hip fracture patient population relative to the changes in the control population.

Inconclusive evidence of potentially delayed care. Medicare spending 30 days after the episode increased by $351 per episode for hip fracture patients who received care at CJR hospitals relative to control hospitals. Based on the recent patient survey, hip fracture patients at CJR hospitals experienced a similar quality of care to hip fracture patients at control hospitals during the episode, and they improved in certain measures of functional status shortly after the episode period relative to control hip fracture patients. Based on these quality findings, we cannot conclude that delayed care or longer recoveries caused the relative increase in post-episode payments for hip fracture patients. We will continue to monitor these results.

Small increase in LEJR volume. For the first time over the course of the CJR Model, we observed a small increase in elective LEJR volume. Patients living in mandatory CJR MSAs in 2022 experienced a relative increase in the number of elective LEJRs of 47 LEJRs per 100,000 Medicare FFS patients. We will continue to monitor changes in LEJR volume.

Conclusion

Since 2016, the CJR Model has successfully reduced Medicare payments for joint replacement procedures while maintaining the quality of care for patients. In PY 6, hospitals participating in the CJR Model achieved significant savings to Medicare through reducing institutional PAC use while improving patient satisfaction and widening access to LEJRs, particularly for patients who are Hispanic and dually eligible.

CJR hospitals continued to focus on reducing institutional PAC use after hospital discharge. To achieve better patient outcomes and reduce costs, health care providers under the CJR Model often optimize care protocols and pathways. These efforts include strategies to reduce patient risk before surgery, standardize surgical techniques, and use evidence-based rehabilitation protocols. Standardized treatment plans can help streamline processes, improve efficiency, and minimize variations in care.

While reducing PAC use, CJR hospitals also continued to maintain the quality of care for LEJR patients. In the most recent survey, patients reported similar levels of functional recovery and

satisfaction with their overall recovery, as well as a need for similar levels of help from their caregivers after returning home. Additionally, patients reported no differences in outcomes among those who had an LEJR in an OP setting compared with those who had an LEJR in an inpatient setting.

New and Emerging Findings

Medicare began covering outpatient knee replacements in 2018. Since that time, inpatient joint replacements have steadily declined, but outpatient procedures have grown faster, resulting in more procedures overall. The COVID-19 pandemic rapidly accelerated this shift as patients avoided hospitalizations and stays in skilled nursing facilities. However, CJR hospitals performed a smaller share of LEJRs in an OP setting than did control hospitals. With less burden and cost for the patient—and the provider—we might expect continued growth in procedures in OP settings. We will continue to monitor this trend in future years.

Hospitals that participated in both a Medicare ACO and the CJR Model indicated that the two programs are viewed as aligned, involve common care redesign strategies, and increase awareness of VBC among hospital employees. Most interviewees viewed the goals of the two programs as similar, stating that both programs share the same dual aims to improve quality of care for patients while reducing expenditures. Hospitals that participated in both VBC initatives employed common strategies, including using patient and population leveldata to inform care pathways, strengthen care coordination, and monitor patient outcomes. Additionally, to ensure alignment with program goals and adopt VBC initiatives, hospitals implemented educational efforts both for hospital staff and other partners such as PAC providers and executive leadership. Participating in VBC programs resulted in a mindset change at hospitals by increasing physicians' awareness of quality throughout the episode, improving understanding of SDOH, and increasing focus on potential economic impacts.

For the first time over the course of the model's implementation, CMS recouped more from hospitals than it paid out. We also observed considerable movement in the amount of NPRA earned by individual hospitals. This change may be due to the new risk adjustment methodology, inclusion of outpatient procedures in the model, or the moving baseline used to set benchmark prices. In PY 6, a high proportion of safety-net CJR hospitals had repayments. A small number of interviews with CJR participants identified as an SNH indicated that they lacked financial success due to low CJR episode volume and low target prices, possibly because the target prices did not reflect the cost of care for their complex patient population.

The CJR Model provides evidence that a mandatory, episode-based payment model that holds providers financially accountable for a well-defined and clinically meaningful episode can achieve significant savings by motivating transformative changes to patient care. We will continue to evaluate potential synergies between ACOs and CJR and further investigate experiences of hospitals who serve diverse patient populations.

I. CJR Model Background

Acronyms[16]	
CBE	Consensus-Based Entity
CFR	Code of Federal Regulations
CJR	Comprehensive Care for Joint Replacement
CMS	Centers for Medicare & Medicaid Services
FFS	fee-for-service
IPPS	Inpatient Prospective Payment System
LEJR	lower extremity joint replacement
MCC	major complications or comorbidities
MSA	metropolitan statistical area
MS-DRG	Medicare Severity-Diagnosis Related Group
NPPGP	non-physician practitioner group practices
OPPS	Outpatient Prospective Payment System
PGP	physician group practice
PY	performance year
Q	quarter
THA	total hip arthroplasty
TKA	total knee arthroplasty

Lower extremity joint replacements (LEJRs) for hips, knees, and ankles represent the most common surgeries Medicare beneficiaries receive. The Centers for Medicare & Medicaid Services (CMS) implemented the Comprehensive Care for Joint Replacement (CJR) Model for LEJRs on April 1, 2016,[17] as part of its strategy to use alternative payment models to slow fee-for-service (FFS) Medicare spending growth by rewarding value rather than volume.

CJR incentivizes hospitals to provide high-quality and cost-effective care. The model requires hospitals in a set of randomly selected metropolitan statistical areas (MSA) to participate and holds them financially accountable for the cost and quality of health care services during and after an LEJR. The CJR Model encourages participant hospitals to coordinate care with the physicians, post-acute care providers, and other providers and clinicians involved in the LEJR throughout the 90-day episodes of care defined by the model rules. Through an annual reconciliation process, participant hospitals may earn additional payments if they achieve cost and quality metrics or face repayments to Medicare if they do not.

CMS originally authorized CJR to run for five performance years (PYs):

- PY 1 covered episodes of care ending in calendar quarter 2 (Q2)[18] 2016 through Q4 2016 (3 quarters)
- PY 2 covered Calendar Year 2017
- PY 3 covered Calendar Year 2018
- PY 4 covered Calendar Year 2019
- PY 5 was extended due to the COVID public health emergency and covered Q1 2020 through Q3 2021 (7 quarters)

In 2021, CMS extended the CJR Model to run for an additional 3 PYs:

- PY 6 covered Q4 2021 through Q4 2022 (5 quarters)
- PY 7 covered Calendar Year 2023

[16] A list of all acronyms used in this report, as well as a glossary of terms, is available in **Appendix A: List of Acronyms and Glossary Terms**.

[17] Centers for Medicare & Medicaid Services. (n.d.). *Comprehensive Care for Joint Replacement Model*. https://www.cms.gov/priorities/innovation/innovation-models/cjr

[18] Calendar year quarters: Q1 = January–March, Q2 = April–June, Q3 = July–September, Q4 = October–December.

- PY 8 covers Calendar Year 2024

This evaluation report focuses on the impact of CJR in PY 6, which includes episodes ending between October 1, 2021, and December 31, 2022. For more information on evaluations of CJR in prior PYs, please refer to the CMS webpage for the CJR Model.[19]

A. Model Design and Definitions

1. Episode Definition

Under the CJR Model, an LEJR episode of care begins with the hospitalization of an eligible Medicare FFS beneficiary at a hospital paid under the Inpatient Prospective Payment System (IPPS) or the Outpatient Prospective Payment System (OPPS) for the surgery and extends through the 90 days after hospital discharge, including the date of discharge. The episode bundle includes related Medicare Part A- and Part B-covered items and services provided during this period, with some exclusions.[20] All providers and suppliers involved in the episode continue to be paid under Medicare's FFS payment system.

Four Medicare Severity-Diagnosis Related Groups (MS-DRGs) identify qualifying surgeries:

1. MS-DRG 469: Major Hip and Knee Joint Replacement or Reattachment of Lower Extremity with Major Complications or Comorbidities (MCC)
2. MS-DRG 470: Major Hip and Knee Joint Replacement or Reattachment of Lower Extremity without MCC
3. MS-DRG 521: Hip Replacement with Principal Diagnosis of Hip Fracture with MCC
4. MS-DRG 522: Hip Replacement with Principal Diagnosis of Hip Fracture without MCC

The CJR Model initially implemented episode-based payments for inpatient LEJRs only. However, CMS removed knee replacements (total knee arthroplasty [TKA]) and hip replacements (total hip arthroplasty [THA]) from Medicare's inpatient-only list effective January 2018 and January 2020, respectively. Subsequently, the CJR Model extended episode-based payments to certain outpatient LEJRs on October 1, 2021, with the start of PY 6.

2. Target Pricing

CMS provides CJR hospitals with preliminary target prices before each PY. The target prices represent the average spending within that hospital's region for each of the four MS-DRGs in the CJR Model based on historical spending data, with a 3% discount applied. The 3% discount serves as Medicare's portion of the savings. CMS then adjusts the target price based on a hospital's

[19] Centers for Medicare & Medicaid Services. (n.d.). *Comprehensive Care for Joint Replacement Model*. https://www.cms.gov/priorities/innovation/innovation-models/cjr

[20] "Unrelated services" are excluded from the episode. These are services for acute clinical conditions that did not arise from existing episode-related chronic clinical conditions or from complications of the LEJR surgery and chronic conditions that are generally not affected by the LEJR procedure or post-surgical care.

composite quality score, a summary score reflecting hospital performance and improvement on two LEJR-related quality measures.[21]

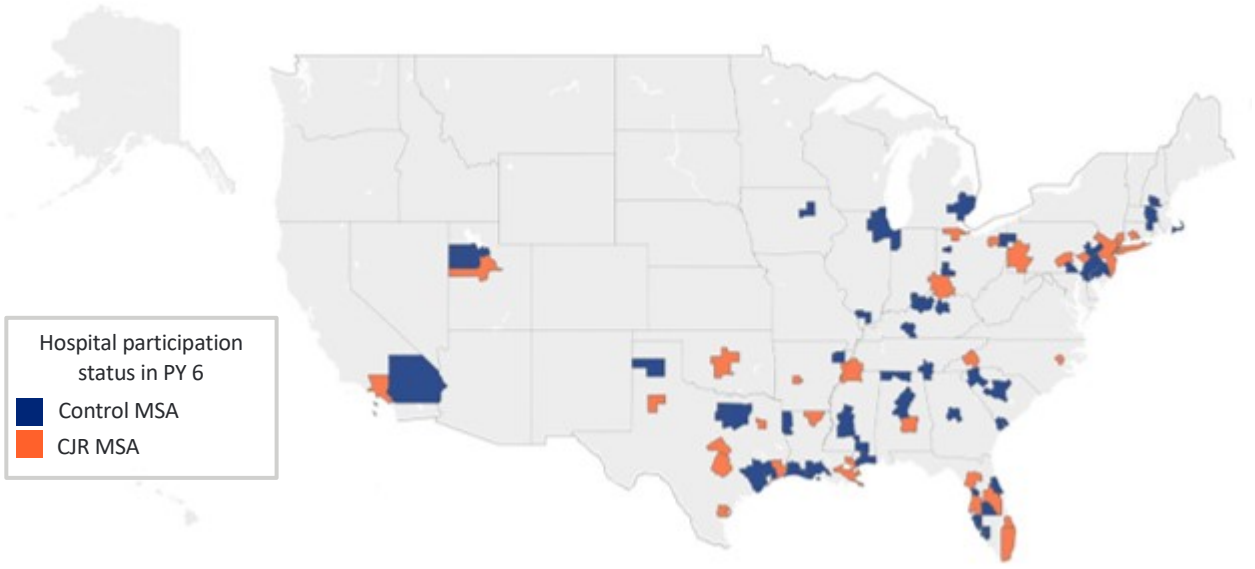
3. Annual Reconciliation

After the end of each model PY, CMS reconciles each participant hospital's LEJR episode payments against the hospital's quality-adjusted target price. For PY 6, CMS based the quality-adjusted target price on a regional average of historical data from 2019. During reconciliation, CMS compares the actual total spending for the episode with the target price for the participant hospital where the beneficiary had the initial LEJR surgery. Depending on the participant hospital's quality and episode spending performance, the hospital may receive an additional payment from Medicare or be required to repay Medicare for a portion of the episode spending.

B. CJR Model Sampling Design and Hospital Participation

CMS requires hospitals in mandatory MSAs to participate in the CJR Model for all Medicare FFS beneficiaries who receive an LEJR categorized in the MS-DRGs included in the model. The mandatory CJR MSAs are primarily in the Middle Atlantic and South Atlantic census divisions, while the control MSAs are primarily in the East North Central and West South Central census divisions (**Exhibit I-1**). **Exhibit B-2** in **Appendix B: Data and Methods** provides a full list of MSAs included in the CJR Model for the CJR and control samples.

Exhibit I-1: The CJR Model Comprises 29 CJR and 41 Control MSAs in PY 6



***Source*:** CJR evaluation team analysis of hospital enrollment data and the CMS Final Rule Medicare Program: Comprehensive Care for Joint Replacement Model Three-Year Extension and Changes to Episode Definition and Pricing, 85 Fed. Reg. 10516 (November 24, 2020) (codified at 42 CFR 510).

Notes: CFR = Code of Federal Regulations; CJR = Comprehensive Care for Joint Replacement; CMS = Centers for Medicare & Medicaid Services; MSA = metropolitan statistical area; PY = performance year.

[21] These two measures are Consensus-Based Entity (CBE) ID 1550: *Hospital-Level Risk-Standardized Complication Rate (RSCR) Following Elective Primary Total Hip Arthroplasty (THA) and/or Total Knee Arthroplasty (TKA)* and CBE ID 0166: *Hospital Consumer Assessment of Healthcare Providers and Systems (HCAHPS) Survey*.

The original mandatory, randomized design of the CJR Model resulted in a diverse group of CJR participant hospitals in 67 randomly selected MSAs, including hospitals that might not voluntarily participate in an episode-based payment model, which allowed a broad test of the CJR Model. For the first 2 PYs, CMS required all acute care hospitals paid under the Medicare IPPS to participate in the model with few exceptions. CMS identified the original 67 mandatory MSAs from 171 MSAs eligible for participation in the model using eight sampling strata based on a median split of MSA population size and quartiles of average MSA historical episode payments. An MSA's probability of selection increased with the payment quartiles to oversample high-payment MSAs for participation in the CJR Model. During model design, CMS hypothesized that higher-payment areas had a greater need and more opportunities for payment reductions. The eligible MSAs that were not selected became a natural control group for evaluating the impact of the CJR Model.

In PY 3, CMS scaled back the CJR Model to the 34 MSAs with the highest historical episode payments (mandatory MSAs). CMS required hospitals in these mandatory MSAs not designated as low volume or rural to continue their participation in the CJR Model. This change reduced the number of hospitals required to participate in the CJR Model from 831 hospitals to 395 hospitals. CMS allowed the remaining hospitals in the 33 lower-payment MSAs (voluntary MSAs) and all hospitals designated as low volume or rural a one-time opportunity to opt in to the CJR Model for PY 3–5. Of the 310 hospitals in the 33 voluntary MSAs, 75 opted to continue their participation in the model.

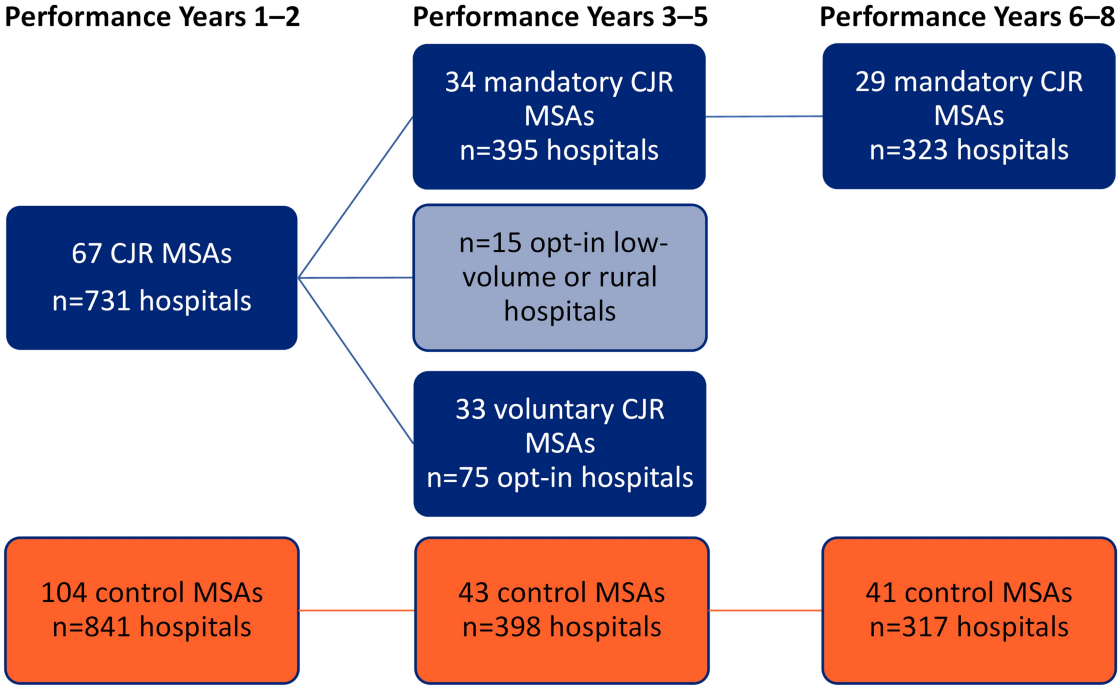
CMS made two main changes to the CJR Model participants for PY 6–8. First, CMS now excludes voluntary opt-in hospitals from the CJR Model. The extension of the CJR Model in 2021 applied only to participant hospitals located in the mandatory MSAs for which participation had been mandatory since the beginning of the model in 2016.

Second, CMS now excludes rural and low-volume hospitals from the CJR Model. This extension also excluded Section 401 rural status hospitals and low-volume hospitals in the mandatory MSAs and any voluntary hospitals in voluntary MSAs that had opted into the model for PY 3–5. The model extension excluded all low-volume or rural hospitals with a CMS Certification Number primary address in the mandatory MSAs. Hospitals approved for rural reclassification no longer participated in the model beginning in PY 6.

Excluding the rural, low-volume, and voluntary opt-in hospitals from the model resulted in a total of 323 CJR hospitals in the 29 mandatory CJR MSAs participating in PY 6–8. This model change removed 72 unique CJR hospitals and five unique CJR MSAs from the sample compared with PY 3–5.[22] The model changes similarly reduced the number of control hospitals from PY 3–5 to PY 6–8. The number of control MSAs decreased from 43 to 41, and the number of control hospitals decreased from 398 to 317 (**Exhibit I-2**).

[22] These five MSAs officially remain CJR mandatory MSAs, but in practice, they contain no eligible hospitals.

Exhibit I-2: In PY 6, 323 Hospitals of the Original 731 Hospitals Remain Required to Participate in the CJR Model After Years of Policy Changes



***Source*:** CJR evaluation team analysis of the *Medicare Program: Comprehensive Care for Joint Replacement Model Three-Year Extension and Changes to Episode Definition and Pricing Performance Year 6 Mid-Year Report.*

Notes: The number of CJR participant hospitals in PY 1–2 was lower than the total number of hospitals chosen for participation because hospitals with no episode volume in the baseline and intervention periods were excluded. The 15 opt-in low-volume or rural hospitals in PY 3–5 were located in mandatory MSAs. CJR = Comprehensive Care for Joint Replacement; MSA = metropolitan statistical area; PY = performance year.

The decision to remove rural and low-volume hospitals from the extension also reduced the variation in size, population dispersion, and rurality within participant hospitals' referral regions between the remaining participant hospitals in PY 6–8. For more information on the CJR population and the effect of policy changes, see **Chapter II: Overview of the CJR Population**.

C. CJR Model Design Changes in PY 6

For the model extension that began in PY 6, CMS implemented multiple changes to the model design.[23]

As noted earlier in this chapter:

- CMS policy changes removed TKA and THA procedures from the inpatient-only list in 2018 and 2020, respectively. CMS began including outpatient TKAs and THAs in the CJR Model in PY 6.

[23] Medicare Program: Comprehensive Care for Joint Replacement Model Three-Year Extension and Changes to Episode Definition and Pricing, 85 Fed. Reg. 10516 (November 24, 2020) (codified at 42 CFR 510).

- The model now includes only hospitals located in the 34 mandatory CJR MSAs and excludes Section 401 rural status, low-volume and voluntary hospitals that had opted into the model for PY 3–5.

CMS revised the methods used to calculate the target prices. CMS implemented these changes with the objectives of making the target prices more accurate, incorporating the inclusion of outpatient TKA/THA episodes, and establishing an "adaptable payment methodology that can sustain adjustments in practice and payment systems over time"[24]:

- The episode target prices for hospitals continue to include a discount factor that adjusts for composite quality score and fracture status, as in prior years. However, the target prices are now also risk-adjusted based on age, dual-eligibility status, and the count of Hierarchical Condition Categories.
- CMS changed the target price calculation from one based on 3 years of claims data to the most recent 1 year of data.
- CMS replaced a national adjustment factor with a retrospective market trend factor that is applied at the reconciliation stage.
- To recognize high-quality care, CMS updated the quality discount factors so that participant hospitals with "good" quality performance receive a 1.5 percentage point reduction to the applicable discount factor and hospitals with "excellent" quality performance receive a 3-percentage point reduction to the applicable discount factor.
- To mitigate the risk of inaccurately capping high-cost cases, CMS also altered the application of spending caps in the calculation.
- Given that all of the previous listed changes were intended to capture the variability in payments more accurately, CMS removed the use of the anchor factor and regional- and hospital-specific anchor weights.

CMS replaced a twice-per-year reconciliation process with the current annual process. This change was designed to reduce the administrative burden for participating hospitals and to improve CMS' ability to account for changes in payment policy and market trends in utilization.

CJR participant hospitals can engage in financial arrangements that allow hospitals to make gainsharing payments to certain providers and collaborators engaged in providing care for LEJR patients. These include Accountable Care Organizations, therapy group practices, physician group practices (PGP), and non-physician practitioner group practices (NPPGP). To align with rules changes for other programs and policies, CMS eliminated the 50% cap on gainsharing payments, distribution payments, and downstream distribution payments when the recipient of these payments are a physician, non-physician practitioner, PGP, or NPPGP.

The beneficiary notification requirement on discharge planning was updated to accommodate the cases where beneficiaries will be discharged the same day following an outpatient procedure.

[24] Medicare Program: Comprehensive Care for Joint Replacement Model Three-Year Extension and Changes to Episode Definition and Pricing, 85 Fed. Reg. 10516 (November 24, 2020) (codified at 42 CFR 510).

D. CJR Participant Care Transformation Strategies

During the first 5 PYs of the CJR Model, hospitals implemented a range of enhanced or new initiatives across the episode of care (before hospitalization, during hospitalization, and after discharge) to decrease the level of acuity of post-acute care use. For example, hospitals offered presurgical joint classes to educate patients, provided physical therapy prior to surgery, and prioritized early identification and intervention with higher-risk patients to optimize outcomes. An accompanying *Care Transformation Report* details hospitals' care transformation strategies.

Hospitals in CJR and control MSAs also transform care in response to incentives from other value-based care programs, including Accountable Care Organizations, Medicare Advantage programs, and contracts with commercial payers. The evaluation's difference-in-differences approach to determining model impacts and the model's design as a mandatory randomized controlled trial should account for the influence of these other value-based care programs.

II. Overview of the CJR Population

In this chapter, we discuss how the overall Comprehensive Care for Joint Replacement (CJR) population changed over time and how it compared with the control population to provide context for findings presented in subsequent chapters. We examined changes to volume in lower extremity joint replacements (LEJRs) for all CJR performance years (PYs) for overall, fracture, and elective episodes in CJR and control hospitals.[25] We also compared outpatient (OP) elective total knee arthroplasty (TKA) and OP elective total hip arthroplasty (THA) rates over time in CJR and control hospitals to examine the response to the changes in OP rules for these procedures. Lastly, we explored descriptive statistics for patients in CJR and control hospitals. We compare patient sociodemographic characteristics, health status, and prior utilization in a post-acute care setting in the baseline period and PY 6 to examine patient population shifts over time.

Acronyms (A–I)	
ACO	Accountable Care Organization
AHA	American Hospital Association
AHRF	Area Health Resources Files
BPCI	Bundled Payments for Care Improvement
CFR	Code of Federal Regulations
CJR	Comprehensive Care for Joint Replacement
CMS	Centers for Medicare & Medicaid Services
DDD	difference-in-difference-in-differences
DiD	difference-in-differences
DSH	Disproportionate Share Hospital
ED	emergency department
ESRD	end-stage renal disease
FFS	fee-for-service
HHA	home health agency
IP	inpatient
IPPS	Inpatient Prospective Payment System
IRF	inpatient rehabilitation facility

A. Summary of Findings

The callout box presents a summary of our findings. Results with an asterisk (*) are key findings.

- (*) More than 70% of elective episodes for both CJR and control hospitals occurred in the OP setting in PY 6. No OP episodes occurred under the CJR Model prior to PY 6, as this is the first PY to reflect the inclusion OP LEJRs as part of the model.
- (*) Differences in CJR and control hospital elective OP LEJR rates continue to persist in PY 6, though OP episode rates increased overall in response to Medicare changes regarding OP episodes.
    - Since the removal of TKAs from the Medicare inpatient-only list in 2018, OP TKAs have increased from 15.0% to 72.4% of all TKAs in CJR hospitals. Similarly, OP TKAs have increased from 25.6% to 78.0% in control hospitals.
    - Since the removal of THAs from the Medicare inpatient-only list in 2020, OP THAs have increased from 31.1% to 71.1% of all THAs in CJR hospitals. Similarly, OP THAs have increased from 36.1% to 78.0% in control hospitals.
- CJR and control patient populations had similar characteristics in PY 6 and in the baseline period, although these characteristics have changed over time.

[25] Includes only mandatory hospitals. Excludes low-volume and rural hospitals in the mandatory CJR MSAs that were not required to continue in the model after PY 2 and hospitals with low volume or Section 401 rural status as of PY 6.

B. Methods

We examined both episode volume and the CJR and control population composition over time. Starting in PY 6, the Centers for Medicare & Medicaid Services (CMS) changed the CJR Model rules to include episodes in a hospital OP setting; therefore, we examined differences in OP LEJR rates between CJR and control hospitals. The share of overall LEJRs occurring in ambulatory surgical centers (ASCs) has slowly but gradually increased, although less than 10% of LEJRs occur in ASCs for both CJR and control metropolitan statistical areas (MSAs).[26] However, CMS does not include OP LEJRs performed at ASCs in the CJR Model, so all measures in this chapter exclude LEJRs at an ASC.

An inpatient LEJR episode begins with an inpatient anchor hospitalization that meets CJR episode eligibility requirements and ends 90 days after discharge.

An OP LEJR episode begins with an anchor procedure that meets CJR OP episode eligibility requirements and ends 90 days after the procedure.

Our analyses do not include ***ASC LEJRs*** as they are excluded from the CJR Model.

1. Episode Volume Over Time

To study trends in the volume of LEJRs, overall as well as separately for fracture and elective episodes at CJR and control hospitals, we calculated and plotted trend charts showing the quarterly count of LEJR episodes between 2016 and 2022 based on episode end date. For PY 1–PY 5, LEJRs were only considered episodes in the CJR Model if they were conducted in the inpatient setting, but OP LEJRs began to be included as episodes in PY 6. To account for the randomized sampling design implemented by CMS when determining CJR MSAs at the beginning of the model, for LEJRs at control hospitals, we weighted the episode counts using the MSA-level sampling weights (refer to **Appendix B: Data and Methods** for details on sampling weights).

2. Proportion of OP LEJR Episodes

To examine the growth in LEJRs performed in an OP setting, we calculated and plotted the proportion of TKAs and elective THAs in an OP setting separately for LEJRs at CJR and control hospitals. We calculated the rates as follows:

- OP TKA rate: Count of episodes for TKAs performed in an OP setting divided by the total number of TKA episodes ending between April 2018 and December 2022
- Elective OP THA rate: Count of episodes for elective THAs performed in an OP setting divided by the total number of elective THA episodes ending between April 2020 and December 2022

For LEJRs at control hospitals, we weighted the episode counts using the MSA-level sampling weights.

[26] Source: CJR evaluation team analysis of Medicare claims and enrollment data for LEJRs between January 2018 and December 2022.

3. Composition of CJR and Control Patient Populations at Baseline and in PY 6

To examine the composition of the CJR and control populations, we calculated the proportion of LEJR episodes for which patients had select patient sociodemographic characteristics, had a diagnosed chronic health condition or comorbidity, or used health care services prior to receiving a LEJR surgery. These proportions were calculated for the baseline period, which included episodes initiated in 2012 through 2014, and for PY 6, which included episodes ending between October 2021 and December 2022.

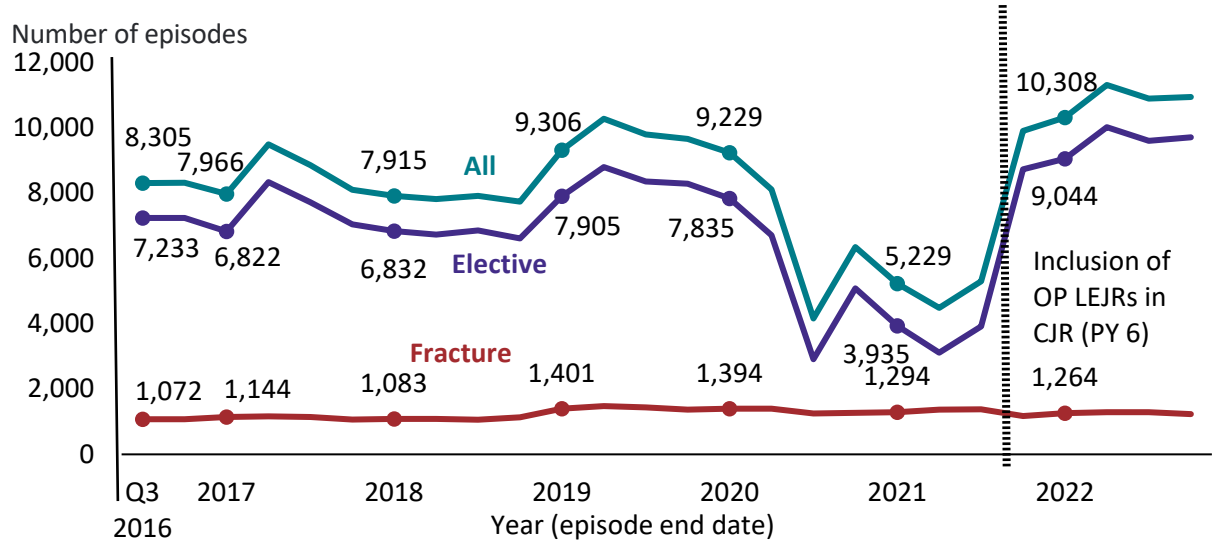
C. Results

1. Volume of LEJR Episodes Over the Course of the CJR Model

Overall, the number of LEJR episodes in the CJR Model and at control hospitals has varied over time, with a slight increase from the beginning of the CJR Model in 2016 to PY 6 in 2022 (**Exhibit II-1** and **II-2**). The number of fracture LEJR episodes remained relatively steady throughout the model, whereas elective episodes fluctuated. These patterns reflect changes in model rules and the broader health care landscape.

CMS removed TKAs and THAs from the inpatient-only list in January 2018 and January 2020, respectively. Prior to PY 6, the CJR Model excluded LEJR procedures in the OP setting, and episode volume under the model declined as OP LEJR procedures gained popularity. As shown in **Exhibit II-1**, the number of elective episodes decreased sharply in PY 5 (2020 and 2021) in conjunction with the removal of THAs from the inpatient-only list and the COVID-19 public health emergency. In PY 6 (2021 and 2022), when CJR began to include OP LEJRs as episodes, the number of elective LEJR episodes increased sharply.

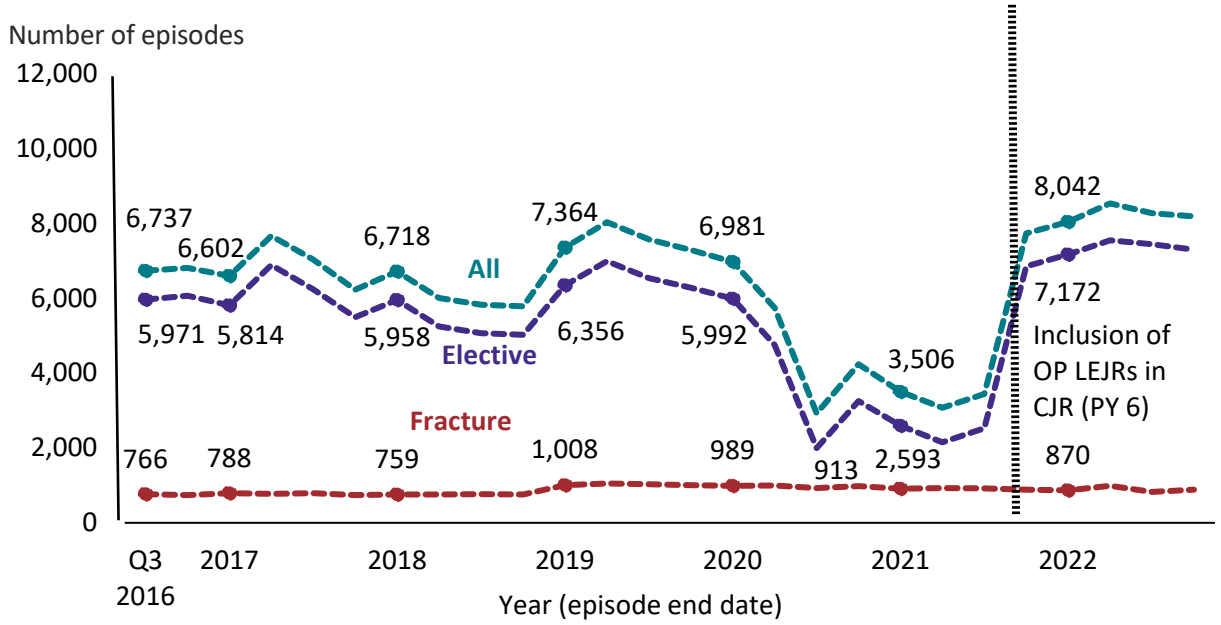
Exhibit II-1: The Number of Episodes Included in the Model at Mandatory CJR Hospitals Has Grown Over Time, With a Sharp Increase in PY 6 (2021 and 2022) Following a Temporary Drop in PY 5 (2020 and 2021)



Source: CJR evaluation team analysis of Medicare claims and enrollment data for LEJR episodes ending between July 2016 (PY 1) and December 2022 (PY 6).

Notes: CJR = Comprehensive Care for Joint Replacement; LEJR = lower extremity joint replacement; OP = outpatient; Q = quarter; PY = performance year.

Exhibit II-2: In Control Hospitals, the Number of Episodes Eligible for the Model Has Also Grown Over Time, With a Sharp Increase in PY 6 (2021 and 2022) Following a Temporary Drop in PY 5 (2020 and 2021)



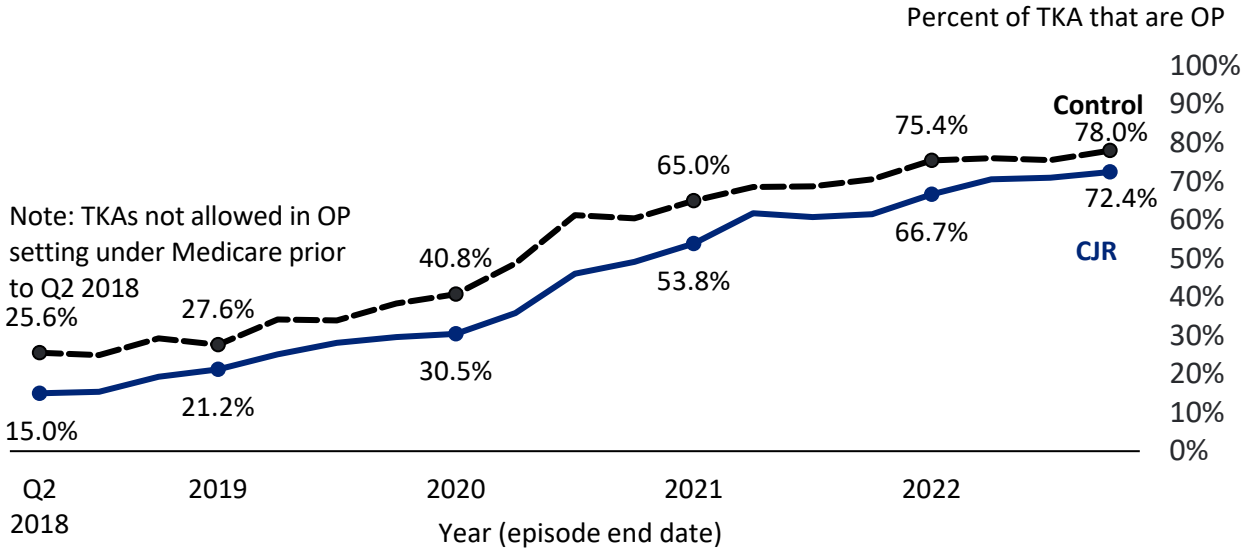
Source: CJR evaluation team analysis of Medicare claims and enrollment data for LEJR episodes ending between July 2016 and December 2022.

Notes: CJR = Comprehensive Care for Joint Replacement; LEJR = lower extremity joint replacement; OP = outpatient; Q = quarter; PY = performance year.

2. Effect of the Outpatient TKA and THA Policy Changes on the CJR Model

Since the inclusion of OP episodes in October 2021, the percentage of OP LEJRs has continued to increase for CJR and control hospitals (**Exhibits II-3 and II-4**). In PY 6, More than 70% of elective episodes for both CJR and control hospitals were in the OP setting. However, a gap remains between CJR and control hospital OP LEJR rates. The share of OP TKAs and THAs in mandatory CJR hospitals was about 6 to 8 percentage points (pp) below the share in control hospitals in PY 6.

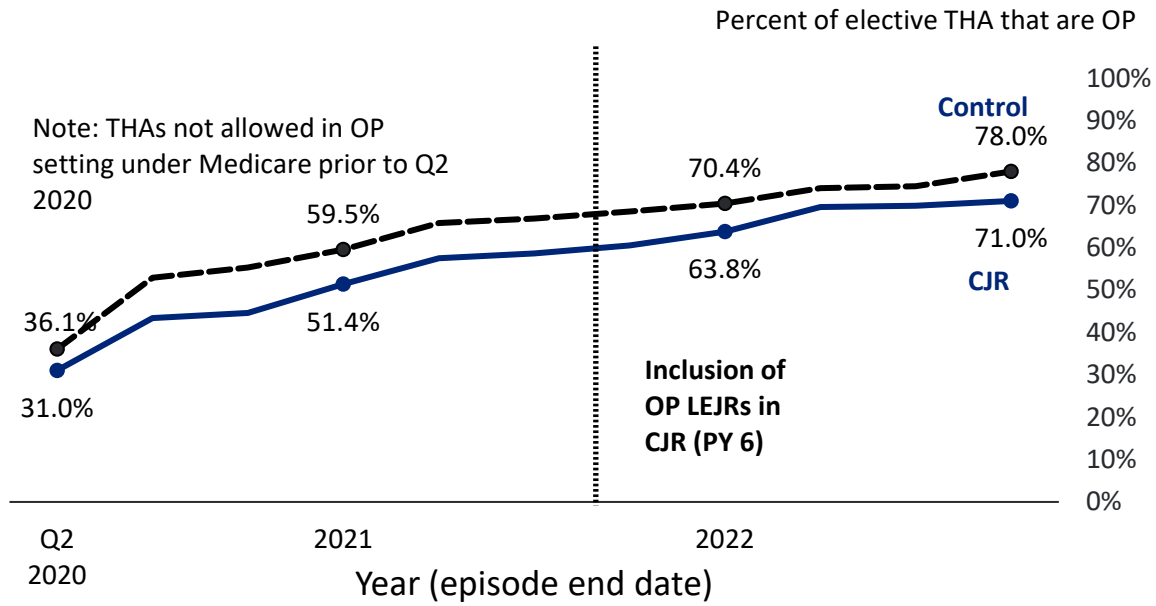
Exhibit II-3: Most TKAs Occurred in the OP Setting, but Mandatory CJR Hospitals Continued to Perform a Smaller Percentage of TKAs in the OP Setting Than Control Hospitals



Source: CJR evaluation team analysis of Medicare claims and enrollment data for LEJR episodes ending between January 2018 and December 2022.

Notes: CJR = Comprehensive Care for Joint Replacement; LEJR = lower extremity joint replacement; OP = outpatient, TKA = total knee arthroplasty; Q = quarter; PY = performance year.

Exhibit II-4: Similar to TKAs, Most Elective THAs Occurred in the OP Setting, but Mandatory CJR Hospitals Continued to Perform a Smaller Percentage of THAs in the OP Setting Than Control Hospitals



Source: CJR evaluation team analysis of Medicare claims and enrollment data for LEJR episodes ending between January 2020 and December 2022.

Notes: CJR = Comprehensive Care for Joint Replacement; LEJR = lower extremity joint replacement; OP = outpatient; THA = total hip arthroplasty; Q = quarter; PY = performance year.

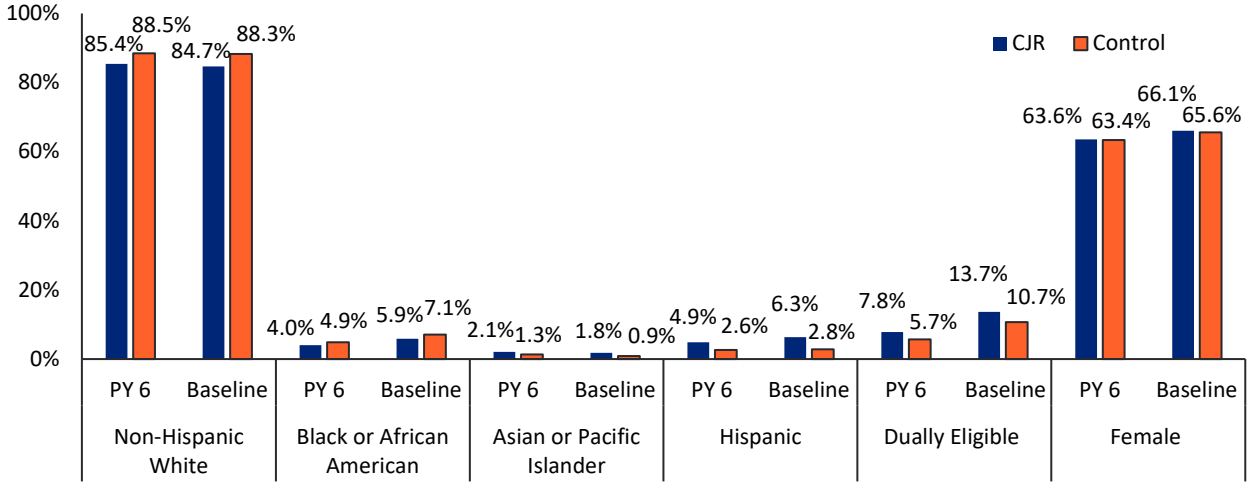
3. CJR and Control Patient Characteristics

The CJR Model was designed to limit patient selection and appropriately account for differences in patient characteristics. In this section, patient population characteristics were examined using descriptive analyses to describe how the patient population compared between CJR and control for PY 6. We present the comparison to baseline as a reference to show how the two populations evolved over time. Overall, patient characteristics remained similar between CJR and control populations (**Exhibits II-5, II-6, and II-7**), although sociodemographic characteristics varied slightly. See **Chapter VI: Analysis of Potential Unintended Consequences of the CJR Model** for additional analyses related to changes in patient characteristics over time for the CJR patient population relative to the control patient population.

a. Sociodemographic Patient Characteristics

For sociodemographic characteristics, we calculated the proportion of episodes for each population (CJR or control) at each measurement period (baseline or PY 6) by sex, race and ethnicity, and enrollment in both Medicare and Medicaid (dual eligibility status). In PY 6, the distribution of CJR patients by race and ethnicity differed somewhat from that of control patients. The CJR population in PY 6 had a higher share of patients who were dually eligible for Medicare and Medicaid than the control population. These distribution trends also existed during the baseline period. The CJR and control population had similar proportions of female patients, although the proportion of female patients have decreased for both populations since the baseline period.

Exhibit II-5: The Distribution of Sociodemographic Patient Characteristics Continued to Differ Slightly Between CJR and Control Populations



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 ending between October 1, 2021, and December 31, 2022 (PY 6 intervention).

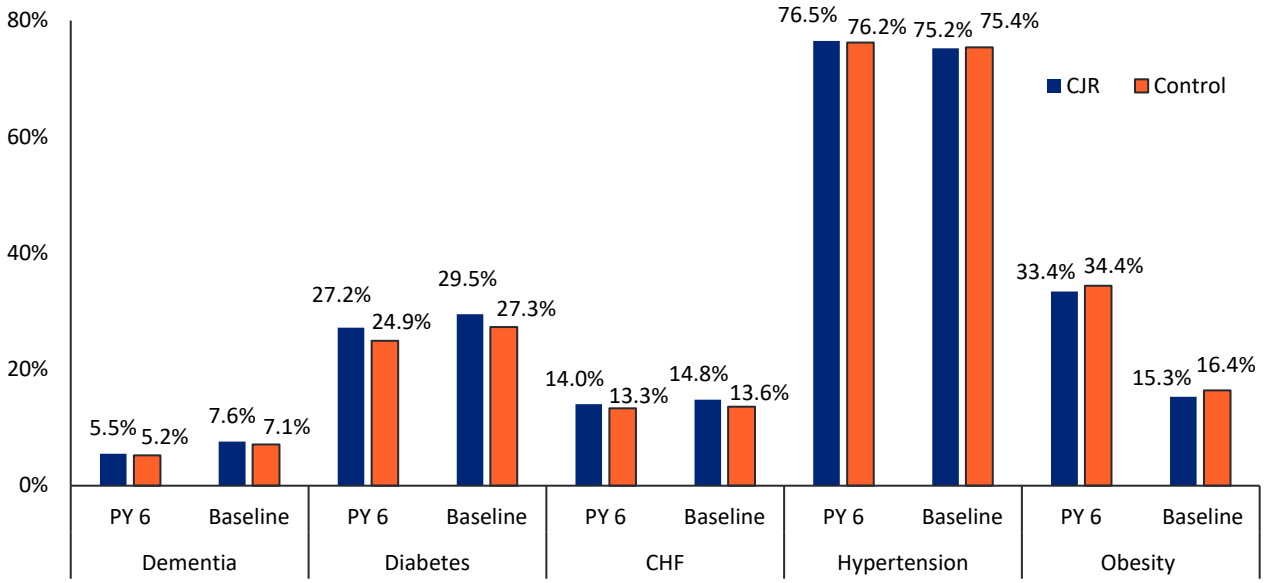
Notes: See **Appendix C: CJR Population Patient Characteristics, Exhibit C-1** for more detailed results. CJR = Comprehensive Care for Joint Replacement; PY = performance year.

b. Chronic Conditions and Comorbidities

For chronic conditions and comorbidities, we calculated the proportion of episodes for each population (CJR or control) at each measurement period (baseline or PY 6) for which a patient had a

diagnosed health condition that may affect their surgical outcomes. These health conditions included dementia, diabetes, congestive heart failure, hypertension, and obesity. We observed one notable difference in these health conditions between the CJR and control patient populations. The CJR population in PY 6 had a larger share of patients with a diabetes diagnosis than the control population by 2.3 pp. We also observed less than a 1 pp difference between the CJR and control population for congestive heart failure, dementia, hypertension, and obesity in PY 6.

Exhibit II-6: Patients in the CJR and Control Populations Had Similar Rates of Chronic Conditions and Comorbidities



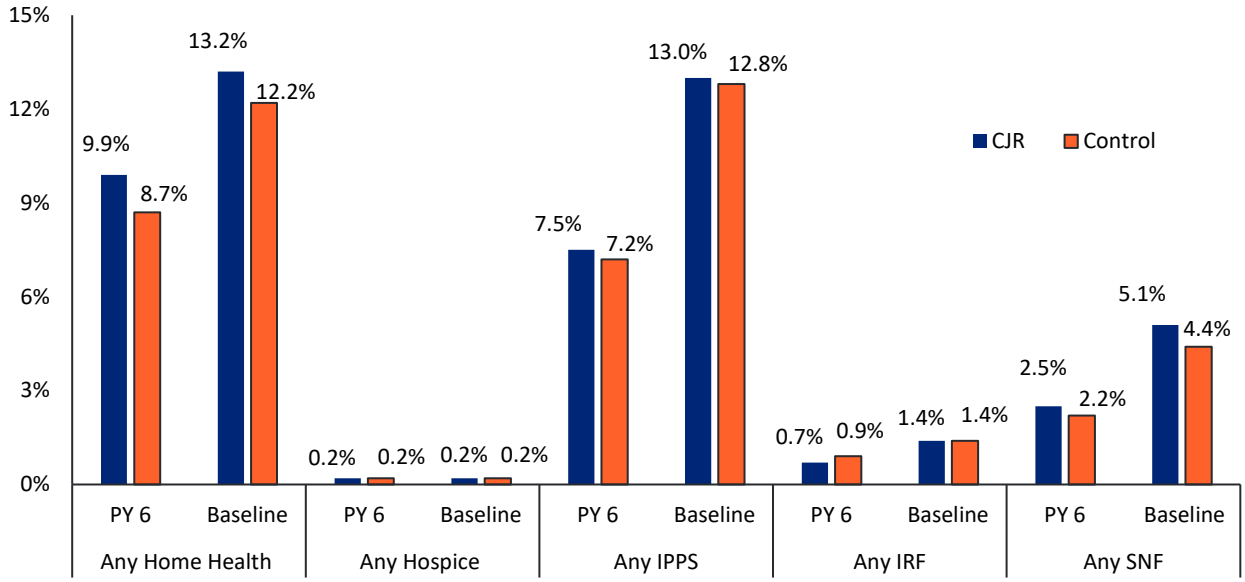
Source: CJR evaluation team analysis of Medicare claims and enrollment data for episodes initiated in 2012 through 2014 that ended between April 2012 and March 2015 (baseline) and episodes ending between October 1, 2021, and December 31, 2022 (PY 6 intervention).

Notes: See **Appendix C: CJR Population Patient Characteristics, Exhibit C-1** for more detailed results. CHF = congestive heart failure; CJR = Comprehensive Care for Joint Replacement; PY = performance year.

c. Care Prior to the LEJR Episode

For prior care utilization, we calculated the proportion of episodes for each population (CJR or control) at each measurement period (baseline or PY 6) for which a patient was admitted to a post-acute care (PAC) setting before their LEJR surgery. For the CJR and control patient populations, we observed decreases in utilization across PAC settings from baseline to PY 6. We observed small differences for CJR compared with control patient populations within 1 pp for all PAC settings except for home health.

Exhibit II-7: Prior Care Decreased Relative to the Baseline Period, Although Utilization Remained Similar Between CJR and Control Patients



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 ending between October 1, 2021, and December 31, 2022 (PY 6 intervention).

Notes: Prior Care was defined as any use during the 6 months prior to the anchor hospitalization for all patients who had an LEJR episode of care at CJR or control hospital. See **Appendix C: CJR Population Patient Characteristics, Exhibit C-1** for more detailed results. CJR = Comprehensive Care for Joint Replacement; IPPS = Inpatient Prospective Payment System; IRF = inpatient rehabilitation facility; PY = performance year; SNF = skilled nursing facility.

III. Impact of the Model

Acronyms	
BPCI	Bundled Payments for Care Improvement
CJR	Comprehensive Care for Joint Replacement
DiD	difference-in-differences
HH	home health
IPPS	Inpatient Prospective Payment System
IRF	inpatient rehabilitation facility
LCI	lower confidence interval
LEJR	lower extremity joint replacement
MS-DRG	Medicare Severity-Diagnosis Related Group
MSA	metropolitan statistical area
OT	occupational therapy
PAC	post-acute care
PT	physical therapy
PY	performance year
SNF	skilled nursing facility
UCI	upper confidence interval

In this chapter, we report findings on the impact of the CJR Model on (A) payments, (B) post-acute care utilization, (C) quality of care, and (D) functional status, pain, and care experience. We calculated payment, utilization, and quality-of-care measures from Medicare claims and enrollment data. We calculated outcome measures related to functional status, pain, and care experience based on responses to a survey of a sample of patients who had a lower extremity joint replacement (LEJR) at Comprehensive Care for Joint Replacement (CJR) or control hospitals in Performance Year 6 (PY 6). We conducted analyses for three population groups: (a) the all LEJR population (elective and hip fracture episodes), (b) the elective-only population, and (c) the hip fracture-only population. As discussed in **Chapter II: Overview of the CJR Population**, most CJR episodes are elective.[27] This chapter is organized by the type of outcome analyzed (A–D above), with subsections for summary of findings, methods, and results. Although the methods for the three claims-based domains are identical, we repeat the description in the text for convenience.

A. What Impact Did the CJR Model Have on Average Episode Payments?

This section addresses how the CJR Model affected average episode payments. Note, the analysis of average episode payments did not incorporate reconciliation payments made to the CJR participants hospitals.[28] We looked at average episode payments, as well as the components that made up the average episode payments. More details on the CJR outcome measures can be found in **Appendix B: Data and Methods, Exhibit B-6**.

1. Summary of Findings

The callout box presents a summary of our findings. Results with an asterisk (*) are key findings.

- (*) CJR hospitals had statistically significant reductions in average episode payments relative to the control hospitals in PY 6, mostly driven by declines in inpatient rehabilitation facility (IRF) payments.
- CJR hospitals had statistically significant reductions in average episode payments relative to the control hospitals for elective LEJRs, mostly driven by declines in IRF payments.
- CJR hospitals did not have statistically significant reductions in average episode payments relative to control hospitals for fracture LEJRs but did have large, offsetting increases in skilled nursing facility (SNF) payments and decreases in IRF payments.

[27] Elective LEJRs are defined as those not caused by hip fractures. More details are in **Appendix B: Data and Methods**.

[28] An analysis of Medicare savings is presented in **Chapter IV: Medicare Program Savings.**

2. Methods

We used a difference-in-differences (DiD) method to estimate the change in average episode payments between the baseline (April 2012 through March 2015) and PY 6 (October 2021 through December 2022) for all LEJR episodes initiated at mandatory CJR hospitals relative to those initiated at control group hospitals.[29] We risk-adjusted estimates for beneficiary, market, and hospital characteristics that can vary over time and between the CJR and control group. We weighted the control group metropolitan statistical areas (MSAs) to be representative of the distribution of the mandatory CJR MSAs. We provide a full description of the methods we used in **Appendix B: Data and Methods**.

3. Results

One of the primary goals of the CJR Model is to reduce average episode payments while maintaining quality of care. Of all the payments included in an episode's bundle of care, payments made towards post-acute care utilization are among the easiest for providers to reduce while still providing all necessary care. As such, changes in post-acute care (PAC) payments, specifically IRF, SNF, and home health (HH), may be likely drivers of change under the CJR Model.

a. All LEJR Episodes

In PY 6, the CJR Model led to an average reduction in average episode payments of $1,012 (p=0.10), driven by a $571 (p=0.06) decrease in IRF payments (**Exhibit III-1**). There were no significant changes in other major components of payments.

We found a $175 (p=0.04) reduction in readmission payments. However, we do not believe it was driven by widespread changes in behavior by CJR hospitals but rather by changes made in a few select CJR hospitals. As reported in the CJR Model evaluation's Fifth Annual Report, during the baseline period, relatively more CJR patients were discharged to Inpatient Prospective Payment System (IPPS) hospitals for rehabilitation.[30] In the intervention period, this practice mostly stopped, resulting in a relative decrease in readmission payments driven by changes in a few CJR participant hospitals. Specifically, 54 CJR participant hospitals in the baseline discharged 1,605 LEJR patients, or 1.4% of the baseline episodes, to IPPS hospitals under Medicare Severity-Diagnosis Related Group (MS-DRG) 945 or 946. In the control group, 13 hospitals discharged less than 0.1% of LEJR patients to IPPS hospitals for rehabilitation during the baseline. This unusual imbalance in the baseline drives our DiD to estimate this reduction in readmission payments in PY 6.

[29] In PY 6, participating CJR hospitals located in the voluntary CJR MSAs and participating low-volume and rural hospitals in the mandatory CJR MSAs were not allowed to continue in the model. We removed these hospitals, and their respective counterparts in the control group, from our sample.

[30] Centers for Medicare & Medicaid Services. (2023). *Comprehensive Care for Joint Replacement Model - Fifth Annual Report*. https://www.cms.gov/priorities/innovation/data-and-reports/2023/cjr-py5-annual-report

Exhibit III-1: During PY 6, for the All-LEJR Population, Declines in IRF Payments Drove a Reduction in Average Episode Payments

Measure	Impact	Impact as a percentage[a]	p-value	90% LCI	90% UCI
Average episode payment	-$1,012	-3.5%	0.10	-$2,017	-$7
SNF payment	-$162	-2.7%	0.60	-$662	$339
IRF payment	-$571	-25.9%	0.06	-$1,067	-$74
HH payment[b]	$190	8.0%	0.34	-$136	$517
Readmission payment	-$175	-14.8%	0.04	-$316	-$34
Anchor payment	$21	0.2%	0.53	-$34	$76
Other Part A	$42	31.5%	0.25	-$19	$102
Other Part B[b]	-$224	-4.5%	0.10	-$450	$1

Source: CJR evaluation team analysis of Medicare claims and enrollment data for episodes initiated between January 1, 2012, and December 31, 2014, that ended between April 1, 2012, and March 31, 2015 (baseline) and episodes that ended between October 1, 2021, and December 31, 2022 (PY 6 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 levels are indicated by red, orange, or yellow shading, respectively. A full table of results, including sample counts and risk-adjusted means, is in **Appendix D: Claims-Based Impact Analyses – Detailed Tables, Exhibit D-6**, and outcome definitions are in **Appendix B: Data and Methods, Exhibit B-6**. CJR = Comprehensive Care for Joint Replacement; DiD = difference-in-differences; HH = home health; IRF = inpatient rehabilitation facility; LCI = lower confidence interval; LEJR = lower extremity joint replacement; PY = performance year; SNF = skilled nursing facility; UCI = upper confidence interval.

[a] Percentages are calculated by dividing the impact by the risk-adjusted mean value of the outcome for CJR episodes in the baseline period.

[b] Indicates that we believe the CJR and control populations may have been on relatively large differential trends in the baseline period for this outcome.

b. Elective Episodes

Overall, the elective-only results were similar to the all-LEJR results, which is unsurprising given that the LEJR population is primarily composed of elective LEJRs.[31] In PY 6, the CJR Model led to a reduction in average episode payments of $1,171 (p=0.09), driven by a $410 (p=0.06) decrease in IRF payments (**Exhibit III-2**). We also found a $256 (p=0.08) reduction in other Part B payments; however, there was substantial evidence that the CJR and control populations were on differential trends in the baseline period, so we did not interpret the estimate as a causal effect of CJR.[32] Although we observed a $164 (p=0.04) decrease in readmission payments, similar to the all-LEJR result for readmission payments, we also do not believe this finding was a causal effect of CJR.

[31] See **Chapter II: Overview of the CJR Population** for more information on the number of elective and fracture LEJRs in the evaluation sample.

[32] See **Appendix D: Claims-Based Impact Analyses – Detailed Tables** for more discussion of parallel trends.

Exhibit III-2: During PY 6, for the Elective LEJR Population, Declines in IRF Payments Drove a Reduction in Average Episode Payments

Measure	Impact	Impact as a percentage[a]	p-value	90% LCI	90% UCI
Average episode payment	-$1,171	-4.5%	0.09	-$2,297	-$44
SNF payment	-$333	-7.9%	0.19	-$754	$89
IRF payment	-$410	-25.1%	0.06	-$763	-$58
HH payment[b]	$211	8.9%	0.33	-$145	$567
Readmission payment	-$164	-17.5%	0.04	-$296	-$31
Anchor payment	$0	0.0%	1.00	-$58	$58
Other Part A	$23	37.8%	0.32	-$15	$61
Other Part B[b]	-$256	-5.4%	0.08	-$499	-$13

Source: CJR evaluation team analysis of Medicare claims and enrollment data for episodes initiated between January 1, 2012, and December 31, 2014, that ended between April 1, 2012, and March 31, 2015 (baseline) and episodes that ended between October 1, 2021, and December 31, 2022 (PY 6 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 shading, respectively. A full table of results, including sample counts and risk-adjusted means, is in **Appendix D: Claims-Based Impact Analyses – Detailed Tables, Exhibit D-7**, and outcome definitions are in **Appendix B: Data and Methods, Exhibit B-6**. CJR = Comprehensive Care for Joint Replacement; DiD = difference-in-differences; HH = home health; IRF = inpatient rehabilitation facility; LCI = lower confidence interval; LEJR = lower extremity joint replacement; PY = performance year; SNF = skilled nursing facility; UCI = upper confidence interval.

[a] Percentages are calculated by dividing the impact by the risk-adjusted mean value of the outcome for CJR episodes in the baseline period.

[b] Indicates that we believe the CJR and control populations may have been on relatively large differential trends in the baseline period for this outcome.

c. *Fracture Episodes*

Unlike in the elective population, the estimate of CJR's impact on average episode payments among fracture episodes was small and not statistically significant, suggesting that the model's effects on elective LEJRs drove the large, statistically significant impact found in the all-LEJR population. Within the fracture population, meaningful changes occurred in some of the payment subcomponents. We estimated a relative $1,575 (p=0.02) *increase* in SNF payments, paired with a relative $1,476 (p=0.05) *decrease* in IRF payments (**Exhibit III-3**), but there was substantial evidence that the CJR and control populations were on differential trends for both of these outcomes in the baseline period. The deviations in trends were such that the true causal impact on SNF payments was likely larger than $1,575, while the true causal impact on IRF payments was likely smaller than –$1,476.[33] Although we are not confident of the exact amounts, we believe there was strong evidence the CJR Model led to a substantial increase in SNF payments and a substantial decrease in IRF payments. Anchor payments also increased by $121 (p=0.06).

[33] See **Appendix D: Claims-Based Impact Analyses – Detailed Tables** for more discussion of parallel trends.

Exhibit III-3: During PY 6, for the Fracture LEJR Population, There Were Large, Offsetting Changes in IRF and SNF Payments, Resulting in No Change for Average Episode Payments

Measure	Impact	Impact as a percentage[a]	p-value	90% LCI	90% UCI
Average episode payment[b]	-$354	-0.8%	0.64	-$1,623	$915
SNF payment[b]	$1,575	9.4%	0.02	$480	$2,671
IRF payment[b]	-$1,476	-27.3%	0.05	-$2,721	-$231
HH payment	-$76	-3.1%	0.33	-$204	$52
Readmission payment	-$179	-7.0%	0.24	-$430	$72
Anchor payment	$121	0.9%	0.06	$16	$225
Other Part A	$101	17.6%	0.35	-$78	$279
Other Part B	-$52	-0.8%	0.67	-$253	$149

Source: CJR evaluation team analysis of Medicare claims and enrollment data for episodes initiated between January 1, 2012, and December 31, 2014, that ended between April 1, 2012, and March 31, 2015 (baseline) and episodes that ended between October 1, 2021, and December 31, 2022 (PY 6 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 shading, respectively. A full table of results, including sample counts and risk-adjusted means, is in **Appendix D: Claims-Based Impact Analyses – Detailed Tables, Exhibit D-8**, and outcome definitions are in **Appendix B: Data and Methods, Exhibit B-6**. CJR = Comprehensive Care for Joint Replacement; DiD = difference-in-differences; HH = home health; IRF = inpatient rehabilitation facility; LCI = lower confidence interval; LEJR = lower extremity joint replacement; PY = performance year; SNF = skilled nursing facility; UCI = upper confidence interval.

[a] Percentages are calculated by dividing the impact by the risk-adjusted mean value of the outcome for CJR episodes in the baseline period.

[b] Indicates that we believe the CJR and control populations may have been on relatively large differential trends in the baseline period for this outcome.

B. What Impact Did the CJR Model Have on Post-Acute Care Use?

This section addresses how the CJR Model affected PAC use for the CJR LEJR population. We analyzed the post-acute care setting to which patients were first discharged, as well as the length of stay for the discharge setting.[34]

1. Summary of Findings

The callout box presents a summary of our findings. Results with an asterisk (*) are key findings.

- (*) CJR hospitals experienced a relative decline in the proportion of patients first discharged to IRFs in the all-LEJR and elective LEJR samples.
- CJR hospitals experienced a relative increase in the proportion of patients first discharged home with home health (HH) in the fracture LEJR sample.

[34] Additional details on all the outcome measures we analyzed are in **Appendix B: Data and Methods**.

2. Methods

We used a DiD method to estimate the change in claims-based measures of utilization between the baseline (April 2012 through March 2015) and PY 6 (October 2021 through 2022) for all LEJR episodes initiated at mandatory CJR hospitals relative to those initiated at control group hospitals.[35] We risk-adjusted estimates for beneficiary, market, and hospital characteristics that can vary over time and between the CJR and control group. We weighted the control group MSAs to be representative of the distribution of the mandatory CJR MSAs.[36]

3. Results

By including the 90 days after the anchor end date in the episode bundle, the CJR Model incentivizes hospitals to both reduce unnecessary utilization and shift utilization from higher cost to lower cost settings, where medically appropriate. Additionally, the model is designed to encourage hospitals and PAC providers to coordinate patient care.[37] One of the drivers of the amount and type of post-acute care utilization is where the patient is first discharged after the anchor stay. As such, we may expect to see a reduction in the use of the most intense PAC discharge destination, IRFs, and an increase in the percentage of patients being discharged home with HH.

a. All LEJR Episodes

The CJR Model led to a decrease of 3.9 percentage points (p=0.05) in the proportion of patients first discharged to an IRF (**Exhibit III-4**).[38] Although we also estimated statistically significant increases in the proportion of patients discharged home with HH and decreases in the proportion discharged home without HH, there was substantial evidence the CJR and control groups were on differential trends in the baseline period for these outcomes. As such, we do not believe our DiD estimates for the proportion of patients discharged home with HH and the proportion discharged home without HH are unbiased causal estimates of the CJR Model.[39]

[35] In PY 6, participating CJR hospitals in the voluntary CJR MSAs and participating low-volume and rural hospitals in the mandatory CJR MSAs were not allowed to continue in the model. We removed these hospitals, and their respective counterparts in the control group, from our sample.

[36] A full description of the methods used are in **Appendix B: Data and Methods**.

[37] Medicare Program: Comprehensive Care for Joint Replacement Model Three-Year Extension and Changes to Episode Definition and Pricing, 85 Fed. Reg. 10516 (November 24, 2020) (codified at 42 CFR 510).

[38] See **Appendix B: Data and Methods, Exhibit B-6** for complete definitions of all outcomes, including the first discharge destination outcomes.

[39] See **Appendix D: Claims-Based Impact Analyses – Detailed Tables** for more discussion of parallel trends.

Exhibit III-4: During PY 6, for the All-LEJR Population, the CJR Model Led to a Large Shift Away From Patients First Being Discharged to IRFs

Measure	Impact	Impact as a percentage[a]	p-value	90% LCI	90% UCI
First PAC SNF[b]	-1.1	-2.6%	0.61	-4.7	2.5
First PAC IRF	-3.9	-28.3%	0.05	-7.2	-0.6
First PAC HH[b]	12.2	34.3%	0.02	3.5	20.9
First PAC home without HH[b]	-7.2	-81.7%	0.07	-13.8	-0.7
SNF days[b]	0.3	1.2%	0.70	-1.1	1.7
IRF days	0.2	2.0%	0.46	-0.3	0.8
HH visits	-0.6	-3.4%	0.37	-1.6	0.5
Any HH use[b]	8.7	11.9%	0.17	-1.8	19.2
Outpatient PT/OT visits	0.1	0.8%	0.84	-0.8	1.0

Source: CJR evaluation team analysis of Medicare claims and enrollment data for episodes initiated between January 1, 2012, and December 31, 2014, that ended between April 1, 2012, and March 31, 2015 (baseline) and episodes that ended between October 1, 2021, and December 31, 2022 (PY 6 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 shading, respectively. A full table of results, including sample counts and risk-adjusted means, is in **Appendix D: Claims-Based Impact Analyses – Detailed Tables, Exhibit D-9**, and outcome definitions are in **Appendix B: Data and Methods, Exhibit B-6**. CJR = Comprehensive Care for Joint Replacement; DiD = difference-in-differences; HH = home health; IRF = inpatient rehabilitation facility; LCI = lower confidence interval; LEJR = lower extremity joint replacement; PAC = post-acute care; PT/OT = physical therapy and occupational therapy; PY = performance year; SNF = skilled nursing facility; UCI = upper confidence interval.

[a] Percentages are calculated by dividing the impact by the risk-adjusted mean value of the outcome for CJR episodes in the baseline period.

[b] Indicates that we believe the CJR and control populations may have been on relatively large differential trends in the baseline period for this outcome.

b. Elective Episodes

Similar to the all-LEJR population, the CJR Model led to a decrease of 3.0 percentage points (p=0.06) in the proportion of patients first discharged to an IRF (**Exhibit III-5**). The average IRF length of stay also increased by 0.5 days (p=0.08) under the CJR Model. We estimated statistically significant changes in the proportion of patients discharged home with HH and the proportion discharged home without HH, but there was substantial evidence the CJR and control groups were on differential trends in the baseline period for these outcomes. Thus, we do not believe these two DiD estimates are an unbiased causal estimate of the CJR Model.[40]

[40] See **Appendix D: Claims-Based Impact Analyses – Detailed Tables** for more discussion of parallel trends.

Exhibit III-5: During PY 6, for the Elective LEJR Population, the CJR Model Led to a Large Shift Away From Patients First Being Discharged to IRFs

Measure	Impact	Impact as a percentage[a]	p-value	90% LCI	90% UCI
First PAC SNF[b]	-2.2	-5.9%	0.34	-6.1	1.6
First PAC IRF	-3.0	-26.2%	0.06	-5.6	-0.4
First PAC HH[b]	13.7	33.7%	0.02	4.3	23.1
First PAC home without HH[b]	-8.5	-87.3%	0.06	-16.0	-1.0
SNF days[b]	-0.5	-2.3%	0.58	-1.9	0.9
IRF days	0.5	4.7%	0.08	0.0	0.9
HH visits	-0.6	-3.9%	0.37	-1.8	0.5
Any HH use	10.3	13.9%	0.16	-1.7	22.2
Outpatient PT/OT visits	0.1	0.6%	0.87	-0.8	0.9

Source: CJR evaluation team analysis of Medicare claims and enrollment data for episodes initiated between January 1, 2012, and December 31, 2014, that ended between April 1, 2012, and March 31, 2015 (baseline) and episodes that ended between October 1, 2021, and December 31, 2022 (PY 6 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 shading, respectively. A full table of results, including sample counts and risk-adjusted means, is in **Appendix D: Claims-Based Impact Analyses – Detailed Tables, Exhibit D-10**, and outcome definitions are in **Appendix B: Data and Methods, Exhibit B-6**. CJR = Comprehensive Care for Joint Replacement; DiD = difference-in-differences; HH = home health; IRF = inpatient rehabilitation facility; LCI = lower confidence interval; LEJR = lower extremity joint replacement; PAC = post-acute care; PT/OT = physical therapy and occupational therapy; PY = performance year; SNF = skilled nursing facility; UCI = upper confidence interval.

[a] Percentages are calculated by dividing the impact by the risk-adjusted mean value of the outcome for CJR episodes in the baseline period.

[b] Indicates that we believe the CJR and control populations may have been on relatively large differential trends in the baseline period for this outcome.

c. *Fracture Episodes*

In PY 6, the proportion of patients first discharged home with HH increased 3.2 percentage points (p=0.04) (**Exhibit III-6**). We also estimated a statistically significant change in the percentage of patients first discharged to an IRF, but there was substantial evidence that the CJR and control groups were on differential trends for both the share of patients first discharged to an IRF and the proportion first discharged to a SNF in the baseline period. Based on the deviations in trends that we observed, the true causal impact on first being discharged to a SNF was likely larger than 3.1 percentage points, while the true causal impact on first being discharged to an IRF was likely smaller than –6.4 percentage points.[41] As such, although we are not confident of the exact amount, we believe there was strong evidence that the CJR Model led to a substantial decrease in the proportion of fracture patients discharged to an IRF.

We estimated a statistically significant increase in the average SNF length of stay, but there was strong evidence that the CJR and control populations were on differential trends in the baseline period for this outcome. We therefore do not interpret the estimate as a causal impact of the CJR Model.[42]

[41] See **Appendix D: Claims-Based Impact Analyses – Detailed Tables** for more discussion of parallel trends.
[42] See **Appendix D: Claims-Based Impact Analyses – Detailed Tables** for more discussion of parallel trends.

Exhibit III-6: During PY 6, for the Fracture LEJR Population, the CJR Model Led to a Large Shift Toward Patients First Being Discharged Home With HH

Measure	Impact	Impact as a percentage[a]	p-value	90% LCI	90% UCI
First PAC SNF[b]	3.1	4.9%	0.18	-0.7	7.0
First PAC IRF[b]	-6.4	-23.5%	0.06	-12.0	-0.7
First PAC HH	3.2	60.9%	0.04	0.6	5.8
First PAC home without HH	0.0	0.9%	0.97	-1.1	1.1
SNF days[b]	1.7	3.9%	0.09	0.0	3.4
IRF days	-0.1	-0.4%	0.88	-0.7	0.5
HH visits	-0.2	-0.8%	0.76	-1.1	0.7
Any HH use	-1.9	-2.7%	0.07	-3.6	-0.2
Outpatient PT/OT visits	0.3	3.1%	0.53	-0.5	1.2

Source: CJR evaluation team analysis of Medicare claims and enrollment data for episodes initiated between January 1, 2012, and December 31, 2014, that ended between April 1, 2012, and March 31, 2015 (baseline) and episodes that ended between October 1, 2021, and December 31, 2022 (PY 6 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 shading, respectively. A full table of results, including sample counts and risk-adjusted means, is in **Appendix D: Claims-Based Impact Analyses – Detailed Tables, Exhibit D-11**, and outcome definitions are in **Appendix B: Data and Methods, Exhibit B-6**. CJR = Comprehensive Care for Joint Replacement; DiD = difference-in-differences; HH = home health; IRF = inpatient rehabilitation facility; LCI = lower confidence interval; LEJR = lower extremity joint replacement; PAC = post-acute care; PT/OT = physical therapy and occupational therapy; PY = performance year; SNF = skilled nursing facility; UCI = upper confidence interval.

[a] Percentages are calculated by dividing the impact by the risk-adjusted mean value of the outcome for CJR episodes in the baseline period.

[b] Indicates that we believe the CJR and control populations may have been on relatively large differential trends in the baseline period for this outcome.

C. What Impact Did the CJR Model Have on Quality of Care?

This section addresses how the CJR Model affected the quality of care for patients with LEJRs. We analyzed the quality of care using claims-based measures: unplanned readmission rates, emergency department use, mortality rates, and complication rates.[43]

1. Summary of Findings

The callout box presents a summary of our findings. Results with an asterisk (*) are key findings.

- (*) The CJR Model maintained quality of care.

2. Methods

We used a DiD method to estimate the change in claims-based quality outcomes between the baseline (April 2012 through March 2015) and PY 6 (October 2021 through December 2022) for all LEJR episodes initiated at mandatory CJR hospitals relative to those initiated at control group

[43] Additional details on all the outcome measures we analyzed are in **Appendix B: Data and Methods**.

hospitals.[44] We risk-adjusted estimates for beneficiary, market, and hospital characteristics that can vary over time and between the CJR and control group. We weighted the control group MSAs to be representative of the distribution of the mandatory CJR MSAs.[45]

3. Results

While much of the CJR Model is focused on reducing average episode payments by better providing the appropriate level of care to patients, especially in the PAC setting, it is crucial that such reductions do not come at the expense of the CJR patients. It is possible CJR hospitals change their behavior in a way that while reducing average episode payments, lowers the quality of care for patients. However, if the changes in care are medically appropriate, we would find that the CJR Model did not lead to changes in quality of care- that is, we should find that the CJR Model maintained the same levels of quality of care.

a. All LEJR Episodes

The CJR Model did not lead to changes in quality of care for any of the four claims-based outcomes assessed in the LEJR population (**Exhibit III-7**).

Exhibit III-7: During PY 6, for the All-LEJR Population, Quality of Care Did Not Change

Measure	Impact	Impact as a percentage[a]	p-value	90% LCI	90% UCI
Unplanned readmission rate	-0.5	4.6%	0.12	-0.9	0.0
Emergency department use rate	0.4	2.8%	0.37	-0.3	1.0
Mortality rate	0.0	-1.8%	0.74	-0.3	0.2
Complication rate	-0.2	-4.8%	0.30	-0.5	0.1

Source: CJR evaluation team analysis of Medicare claims and enrollment data for episodes initiated between January 1, 2012, and December 31, 2014, that ended between April 1, 2012, and March 31, 2015 (baseline) and episodes that ended between October 1, 2021, and December 31, 2022 (PY 6 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 shading, respectively. A full table of results, including sample counts and risk-adjusted means, is in **Appendix D: Claims-Based Impact Analyses – Detailed Tables, Exhibit D-12**, and outcome definitions are in **Appendix B: Data and Methods, Exhibit B-6**. CJR = Comprehensive Care for Joint Replacement; DiD = difference-in-differences; LCI = lower confidence interval; LEJR = lower extremity joint replacement; PY = performance year; UCI = upper confidence interval.

[a] Percentages are calculated by dividing the impact by the risk-adjusted mean value of the outcome for CJR episodes in the baseline period.

b. Elective Episodes

The CJR Model did not lead to changes in quality of care for any of the four outcomes assessed in the elective population (**Exhibit III-8**).

[44] In PY 6, participating CJR hospitals in the voluntary CJR MSAs and participating low-volume and rural hospitals in the mandatory CJR MSAs were not allowed to continue in the model. We removed these hospitals, and their respective counterparts in the control group, from our sample.

[45] A full description of the methods used are in **Appendix B: Data and Method**.

Exhibit III-8: During PY 6, for the Elective LEJR Population, Quality of Care Did Not Change

Measure	Impact	Impact as a percentage[a]	p-value	90% LCI	90% UCI
Unplanned readmission rate	-0.3	-3.8%	0.22	-0.7	0.1
Emergency department use rate	0.5	4.5%	0.22	-0.2	1.3
Mortality rate	0.2	-4.3%	0.70	-0.1	0.1
Complication rate	-0.2	-7.3%	0.16	-0.5	0.0

Source: CJR evaluation team analysis of Medicare claims and enrollment data for episodes initiated between January 1, 2012, and December 31, 2014, that ended between April 1, 2012, and March 31, 2015 (baseline) and episodes that ended between October 1, 2021, and December 31, 2022 (PY 6 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 shading, respectively. A full table of results, including sample counts and risk-adjusted means, is in **Appendix D: Claims-Based Impact Analyses – Detailed Tables, Exhibit D-13**, and outcome definitions are in **Appendix B: Data and Methods, Exhibit B-6**. CJR = Comprehensive Care for Joint Replacement; DiD = difference-in-differences; LCI = lower confidence interval; LEJR = lower extremity joint replacement; PY = performance year; UCI = upper confidence interval.

[a] Percentages are calculated by dividing the impact by the risk-adjusted mean value of the outcome for CJR episodes in the baseline period.

c. Fracture Episodes

The CJR Model did not lead to changes in quality of care for any of the four outcomes assessed in the fracture population (**Exhibit III-9**).

Exhibit III-9: During PY 6, for the Fracture LEJR Population, Quality of Care Did Not Change

Measure	Impact	Impact as a percentage[a]	p-value	90% LCI	90% UCI
Unplanned readmission rate	-1.3	-6.1%	0.19	-2.9	0.3
Emergency department use rate	-0.6	-3.4%	0.49	-2.0	0.8
Mortality rate	-0.1	-0.9%	0.87	-1.3	1.0
Complication rate	0.1	1.0%	0.87	-0.9	1.1

Source: CJR evaluation team analysis of Medicare claims and enrollment data for episodes initiated between January 1, 2012, and December 31, 2014, that ended between April 1, 2012, and March 31, 2015 (baseline) and episodes that ended between October 1, 2021, and December 31, 2022 (PY 6 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 shading, respectively. A full table of results, including sample counts and risk-adjusted means, is in **Appendix D: Claims-Based Impact Analyses –Detailed Tables, Exhibit D-14**, and outcome definitions are in **Appendix B: Data and Methods, Exhibit B-6**. CJR = Comprehensive Care for Joint Replacement; DiD = difference-in-differences; LCI = lower confidence interval; LEJR = lower extremity joint replacement; PY = performance year; UCI = upper confidence interval.

[a] Percentages are calculated by dividing the impact by the risk-adjusted mean value of the outcome for CJR episodes in the baseline period.

D. What Impact Did the CJR Model Have on Functional Status and Care Experiences?

We collected information from a sample of patients who had an LEJR at a CJR or control hospital after the end of their episodes to learn about their experiences. The patient survey measures clinical

outcomes and potential unintended consequences from the patient's perspective that cannot be captured by claims or other secondary data sources. This section presents results for patient-reported outcomes related to functional status, satisfaction with care, experience with care transitions, and caregiver help from our patient survey.

1. Summary of Findings

The callout box presents a summary of our findings. Results with an asterisk (*) are key findings.

- (*) The CJR Model did not lead to systematic changes in functional status or care experiences.
- Across all LEJR episodes (elective inpatient, elective outpatient, and hip fractures), CJR and control survey respondents with inpatient discharges or outpatient procedures from July through October 2022 had similar self-reported changes in functional status and pain, had similar satisfaction with their overall recovery, and required similar levels of help from their caregivers after returning home, on average.
- Among the subset of respondents with hip fracture episodes, CJR respondents reported significantly lower levels of decline in function from before their fracture through the end of their episode than did control respondents for measures of using stairs and dependence on a mobility aid. CJR respondents also reported higher satisfaction with recovery and care management than control respondents and were more likely to agree that they were discharged from the hospital at the right time.
- There were no systematic differences in CJR Model impacts on patient-reported outcomes between elective inpatient and elective outpatient episodes.

2. Methods

We surveyed beneficiaries after the end of their LEJR episode to determine whether CJR patients differed from control patients on several patient-reported outcomes.[46] Measures included 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. This analysis employed cross-sectional comparisons of patient-reported outcomes instead of a DiD method utilized for our claims-based outcomes, since we did not have survey data prior to the start of the CJR Model. We estimated risk-adjusted differences between CJR and control respondents, accounting for beneficiary, episode, hospital, and MSA attributes. This analysis employs cross-sectional comparisons of patient-reported outcomes. We provide additional detail on regression specifications in **Appendix B: Data and Methods**.

The survey had eight questions about dimensions of function and pain scale, including difficulty with rising, standing, walking, using stairs, and toileting; using a mobility aid (for example, a cane or walker); pain interfering with daily activities; and use of pain medication. Responses were

[46] The mean time at which surveys were returned was 127 days after hospital discharge or outpatient procedure (37 days after the conclusion of the patient's 90-day post-discharge period). The difference in average response time between CJR and control respondents was less than one day (p=0.30).

collected on a Likert scale. Most measures had five response options, including "extreme," "severe," "moderate," "mild," and "none." For each measure, respondents were asked to recall their status in the week before surgery and to indicate their status at the time of the survey. From these data, we constructed a measure of change in function or pain from prior to surgery through the time of the survey.[47] Because differences in levels of function are not an intuitive measure, we express changes as a percentage of average reported prehospital function.[48]

The survey also had four measures of caregiver help and six measures of satisfaction. To facilitate interpretation, we scaled these measures from 0 (worst outcome) to 100 (best outcome) points. We reported measures of experience with care transitions in percentage terms (with differences interpreted on a percentage point scale) because these reflect binary measures. We scaled all outcomes in all domains such that 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.

Beginning in PY 6, we surveyed patients with all types of LEJR episodes: elective inpatient (25.9% of LEJRs during our sampling period), elective outpatient (61.8% of LEJRs), and fracture episodes (12.3% of LEJRs). Prior surveys only included patients with inpatient LEJRs. We collected data in two batches to achieve the desired sample size. The first batch covered LEJR episodes with inpatient discharges or outpatient procedures in July or August 2022 during PY 6. The second batch covered LEJR episodes with inpatient discharges or outpatient procedures discharged in September or October 2022. We report starting samples and response rates in **Appendix B: Data and Methods, Exhibit B-12**.

3. Results

a. All LEJR Patients

This section presents the patient survey results related to functional status and pain, satisfaction with overall recovery and care management, experience with care transitions, and caregiver help needed after returning home for all LEJR patients.

Functional Status and Pain

Respondents from mandatory CJR hospitals and control respondents reported improvements from before their surgery to after the episode on all eight measures of functional status and pain, and improvements were similar in magnitude (**Exhibit III-10**). Improvements were highest for pain interfering with normal activities and smallest for use of a mobility aid. Differences between CJR and control respondents across the measures varied in direction and were not statistically significant.

[47] The range of each change measure depends on the number of response options. For example, a measure with five categories of function can decline 4 levels (from "no difficulty" to "extreme difficulty") or improve 4 levels (from "extreme difficulty" to "no difficulty") or anywhere in between.

[48] For example, if the average respondent scored 4 out of 5 on prehospital function, and the estimated difference was 0.4 levels, then this would translate to a 10% difference between the CJR and control groups (0.4/4.0 = 0.10).

Exhibit III-10: Among All LEJR Patients, CJR and Control Survey Respondents Experienced Similar Improvement in Functional Status and Pain

Measure	Range	Number of respondents		Risk-adjusted average of change before and after LEJR		Estimated difference between CJR and control	p-value
		CJR	Control	CJR	Control		
Ability to walk by yourself without resting	-4 to 4	4,697	4,501	0.8	0.8	0.0 (1.6%)	0.13
Difficulty walking up or down 12 stairs	-3 to 3	4,530	4,327	0.8	0.8	0.0 (0.5%)	0.66
Difficulty rising from sitting	-4 to 4	4,782	4,575	1.2	1.2	0.0 (0.1%)	0.91
Difficulty standing	-4 to 4	4,788	4,579	1.2	1.2	0.0 (0.5%)	0.39
Use of a mobility aid	-2 to 2	4,762	4,556	0.2	0.1	0.0 (0.9%)	0.34
Difficulty getting on/off the toilet	-4 to 4	4,792	4,577	1.3	1.3	0.0 (-0.4%)	0.47
Frequency that pain interferes with normal activities	-4 to 4	4,778	4,576	2.0	2.0	0.0 (-1.1%)	0.27
Medication use for pain in the joint you had replaced	-3 to 3	4,698	4,491	0.6	0.6	0.0 (0.0%)	0.96

Source: CJR evaluation team analysis of patient survey data for episodes with inpatient discharge or outpatient procedure in July – October 2022.

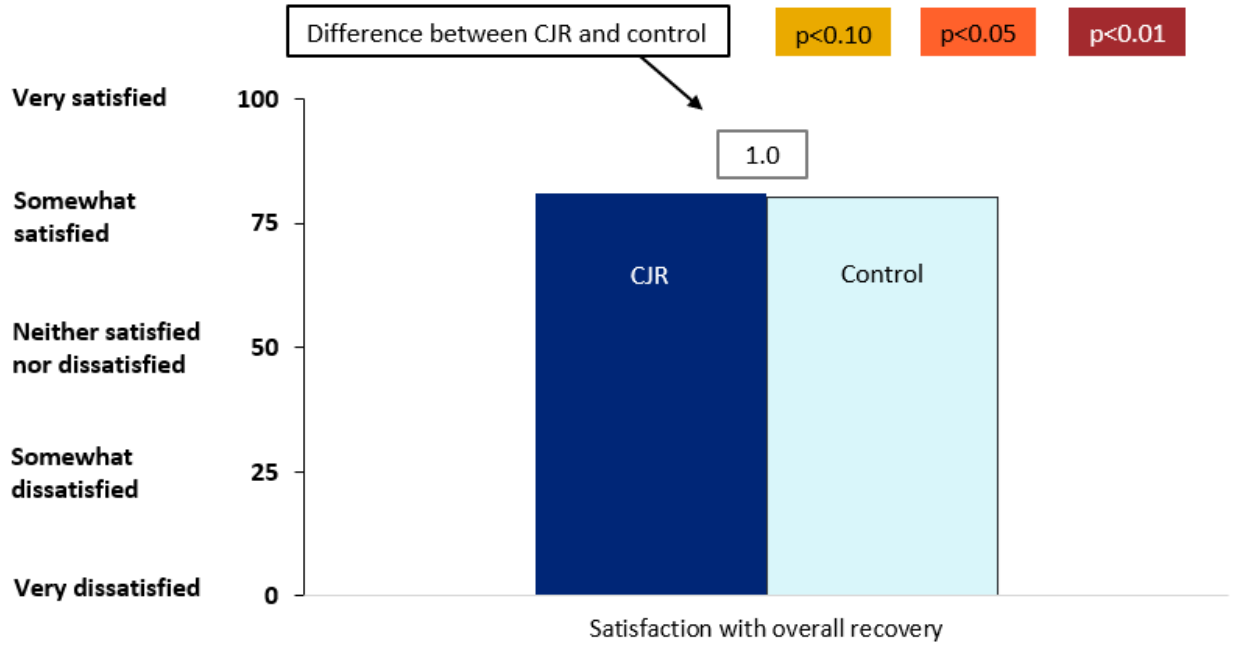
Notes: The estimates in this exhibit are the results of a cross-sectional regression model, weighted for sampling and nonresponse. The change in each measure refers to the difference between a respondent's self-reported status at the time of the survey and the respondent's recalled status prior to the procedure. Estimated changes, and the difference of changes between 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. See **Appendix B: Data and Methods, Exhibit B-15** for the list of covariates included in the regression and weighting procedures. Estimates that are significant at the 1%, 5%, or 10% significance level are indicated by red, orange, or yellow shading, respectively. CJR = Comprehensive Care for Joint Replacement; LEJR = lower extremity joint replacement.

Satisfaction

More than 60% of respondents in both groups were very satisfied with their overall recovery, while fewer than 15% reported any level of dissatisfaction. For CJR and control respondents' average satisfaction with recovery was 81 and 80, respectively, on a 100-point scale, and differences between respondent groups were less than 1 percentage point (**Exhibit III-11**).

Similarly, average satisfaction with all four measures of care management ranged from 77 to 84 on a 100-point scale in both groups. Differences were less than 1.5 points and not statistically significant (results presented in **Appendix E: Patient Survey Results, Exhibit E-5**). For each of the four measures, roughly 65–75 of 100 respondents were very satisfied, while fewer than 20% reported any level of dissatisfaction (**Exhibit III-12**).

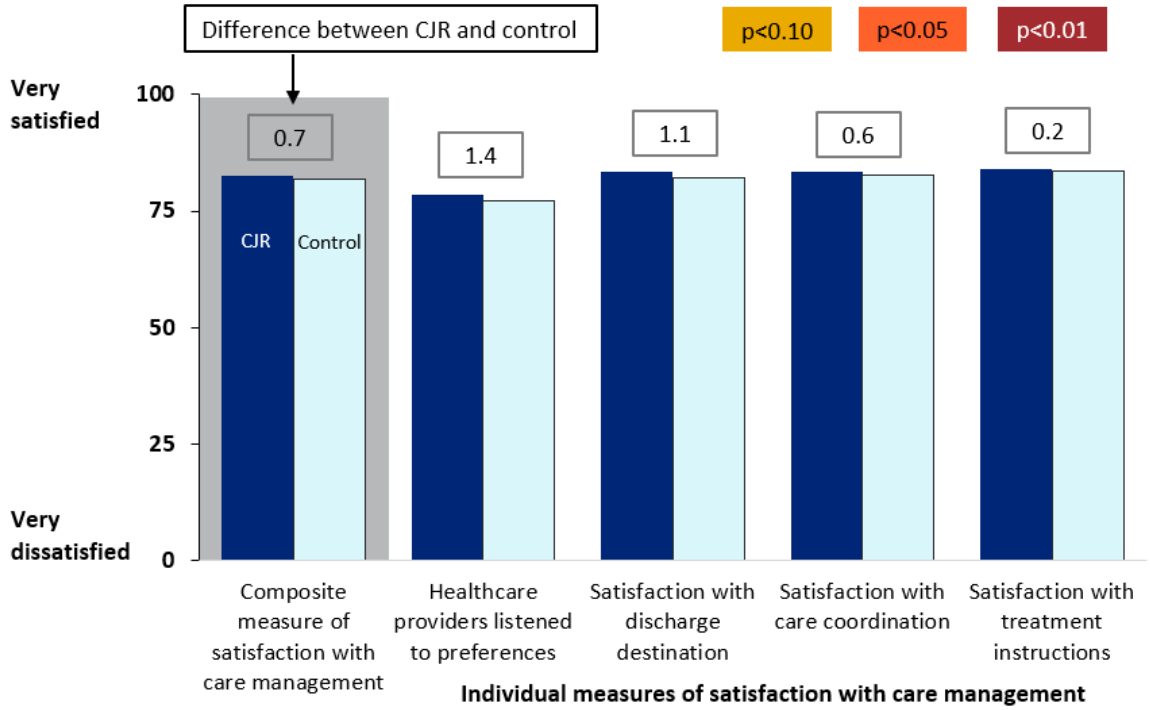
Exhibit III-11: Among All LEJR Patients, CJR and Control Survey Respondents Reported Similar Satisfaction With Overall Recovery



Source: CJR evaluation team analysis of patient survey data for episodes with inpatient discharge or outpatient procedure in July – October 2022.

Notes: The vertical bars depict the risk-adjusted average for CJR and control respondents. The difference between CJR and control is reported at the top of the figure in a box. Measures of satisfaction consisted of a Likert scale with five levels and were normalized so that the lowest response category ("very dissatisfied") yielded a score of 0 and the highest response category ("very satisfied") yielded a score of 100. The estimates in this exhibit are the results of a cross-sectional regression model, weighted for sampling and nonresponse. See **Appendix B: Data and Methods, Exhibit B-15** for the list of covariates included in the regression and weighting procedures. Estimates that are significant at the 1%, 5%, or 10% significance level are indicated by red, orange, or yellow shading, respectively. CJR = Comprehensive Care for Joint Replacement; LEJR = lower extremity joint replacement.

Exhibit III-12: Among All LEJR Patients, CJR and Control Survey Respondents Reported Similar Satisfaction With Care Management



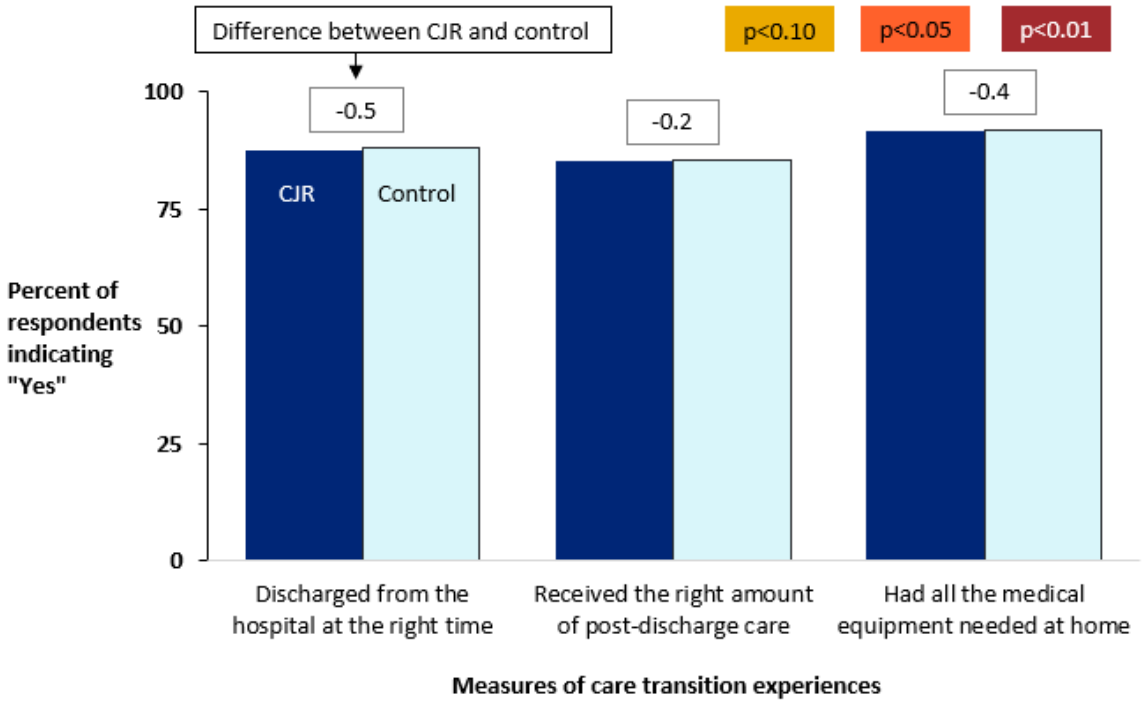
Source: CJR evaluation team analysis of patient survey data for episodes with inpatient discharge or outpatient procedure in July – October 2022.

Notes: The vertical bars depict the risk-adjusted average for CJR and control respondents. The difference between CJR and control is reported at the top of the figure in boxes. Measures of satisfaction consisted of a Likert scale with five levels and were normalized so that the lowest response category ("very dissatisfied") yielded a score of 0 and the highest response category ("very satisfied") yielded a score of 100. The estimates in this exhibit are the results of a cross-sectional regression model, weighted for sampling and nonresponse. See **Appendix B: Data and Methods, Exhibit B-15** for the list of covariates included in the regression and weighting procedures. Estimates that are significant at the 1%, 5%, or 10% significance level are indicated by red, orange, or yellow shaded cells, respectively. CJR = Comprehensive Care for Joint Replacement; LEJR = lower extremity joint replacement.

Experience With Care Transitions

More than 85% of respondents in the CJR and control groups agreed that they were discharged from the hospital at the right time, and a similar proportion agreed that they received the right amount of care after discharge. More than 90% of respondents in both groups agreed that they had all the medical equipment they needed at home (**Exhibit III-13**). The differences were less than half a percentage point and not statistically significant.

Exhibit III-13: Among All LEJR Patients, Care Transitions Were Similarly Positive Between CJR and Control Respondents



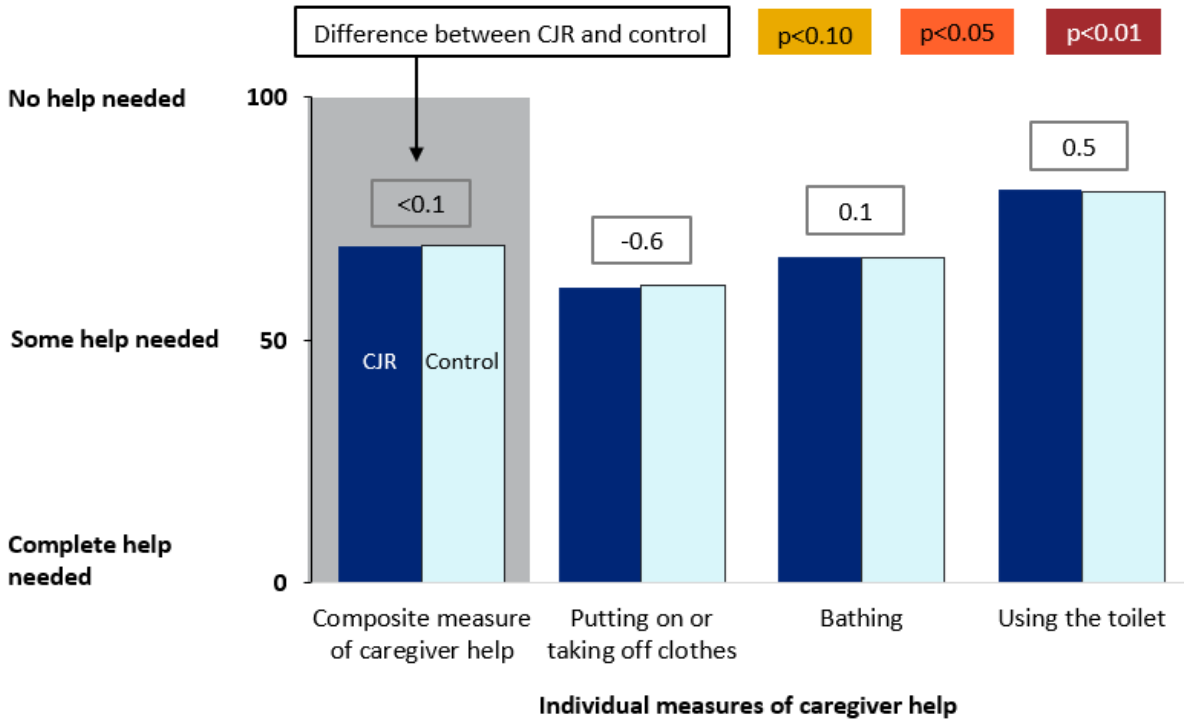
Source: CJR evaluation team analysis of patient survey data for episodes with inpatient discharge or outpatient procedure in July – October 2022.

Notes: The vertical bars depict the risk-adjusted average for CJR and control respondents. The difference between CJR and control is reported at the top of the figure in boxes. The estimates in this exhibit are the results of a cross-sectional regression model, weighted for sampling and nonresponse. See **Appendix B: Data and Methods, Exhibit B-15** for the list of covariates included in the regression and weighting procedures. Estimates that are significant at the 1%, 5%, or 10% significance level are indicated by red, orange, or yellow shading, respectively. CJR = Comprehensive Care for Joint Replacement; LEJR = lower extremity joint replacement.

Caregiver Help

About 96% of both CJR and control respondents received help from a caregiver after returning home. Among those who received help from a caregiver, most respondents in both the CJR and control group needed at least some help with dressing and bathing, but respondents needed less help using the toilet. Only about 1 in 10 respondents receiving caregiver help reported "complete help needed" with any measure. Most variation across measures was between "some help needed" and "no help needed." Overall, CJR and control respondents reported needing similar amounts of help with putting on and taking off clothes, bathing, and using the toilet (**Exhibit III-14**).

Exhibit III-14: Among All LEJR Patients, CJR and Control Respondents Needed Similar Amounts of Caregiver Help at Home



Source: CJR evaluation team analysis of patient survey data for episodes with inpatient discharge or outpatient procedure in July – October 2022.

Notes: The vertical bars depict the risk-adjusted average for CJR and control respondents. The difference between CJR and control is reported at the top of the figure in boxes. Measures consisted of a Likert scale with five levels and were normalized so that the lowest response category ("complete help needed") yielded a score of 0 and the highest response category ("no help needed") yielded a score of 100. The estimates in this exhibit are the results of a cross-sectional regression model, weighted for sampling and nonresponse. See **Appendix B: Data and Methods, Exhibit B-15** for the list of covariates included in the regression and weighting procedures. Estimates that are significant at the 1%, 5%, or 10% significance level are indicated by red, orange, or yellow shading, respectively. CJR = Comprehensive Care for Joint Replacement; LEJR = lower extremity joint replacement.

b. *Elective Patients*

We separately analyzed outcomes for respondents with elective inpatient LEJR episodes and elective outpatient LEJR episodes and then compared CJR Model impacts across the two settings. Our results did not indicate any pattern of differences between CJR and control respondents in either LEJR setting for any of the survey domains. CJR Model impacts were similarly small, and mostly insignificant, for all elective episodes, regardless of setting. We present results in **Appendix E: Patient Survey Results**.

c. *Fracture Patients*

This section presents the patient survey results related to functional status and pain, satisfaction with overall recovery and care management, experience with care transitions, and caregiver help needed after returning home across survey respondents with hip fracture surgeries.

Functional Status and Pain

Due to the urgent and unexpected nature of hip fractures, we expect fracture patients to rate their functional status lower after their episode than before the fracture requiring their LEJR. Both CJR and control patients reported a decline in function from before their hospitalization to after the end of the episode. However, CJR respondents with hip fractures reported lower levels of decline than control respondents with hip fractures in difficulty walking up or down stairs (5.6%; p=0.05) and dependence on a mobility aid (4.4%; p=0.02) (**Exhibit III-15**).[49] These differences are roughly equal to 4 to 5 additional CJR respondents out of 100 reporting better functional status relative to control respondents.

Exhibit III-15: CJR Respondents With Hip Fractures Reported Less of a Decline Than Control Respondents in Walking Up or Down Stairs and Dependence on a Mobility Device

Measure	Range	Number of respondents		Risk-adjusted average of change before and after LEJR		Estimated difference between CJR and control	p-value
		CJR	Control	CJR	Control		
Ability to walk by yourself without resting	-4 to 4	512	582	-0.7	-0.8	0.1 (2.1%)	0.34
Difficulty walking up or down 12 stairs	-3 to 3	475	548	-0.4	-0.6	0.2 (5.6%)	0.05
Difficulty rising from sitting	-4 to 4	529	607	-0.3	-0.3	0.0 (0.0%)	1.00
Difficulty standing	-4 to 4	532	602	-0.3	-0.2	0.0 (-1.2%)	0.46
Use of a mobility aid	-2 to 2	523	603	-0.6	-0.7	0.1 (4.4%)	0.02
Difficulty getting on/off the toilet	-4 to 4	527	606	0.0	-0.1	0.0 (0.5%)	0.72
Frequency that pain interferes with normal activities	-4 to 4	521	598	-0.3	-0.3	0.0 (-0.7%)	0.67
Medication use for pain in the joint you had replaced	-3 to 3	518	586	-0.2	-0.2	0.0 (0.8%)	0.59

Source: CJR evaluation team analysis of patient survey data for episodes with inpatient discharge or outpatient procedure in July–October 2022.

Notes: The estimates in this exhibit are the results of a cross-sectional regression model, weighted for sampling and nonresponse. The change in each 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 procedure. Estimated changes, and the difference of changes between 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. See **Appendix B: Data and Methods, Exhibit B-15** for the list of covariates included in the regression and weighting procedures. Estimates that are significant at the 1%, 5%, or 10% significance level are indicated by red, orange, or yellow shading, respectively. CJR = Comprehensive Care for Joint Replacement.

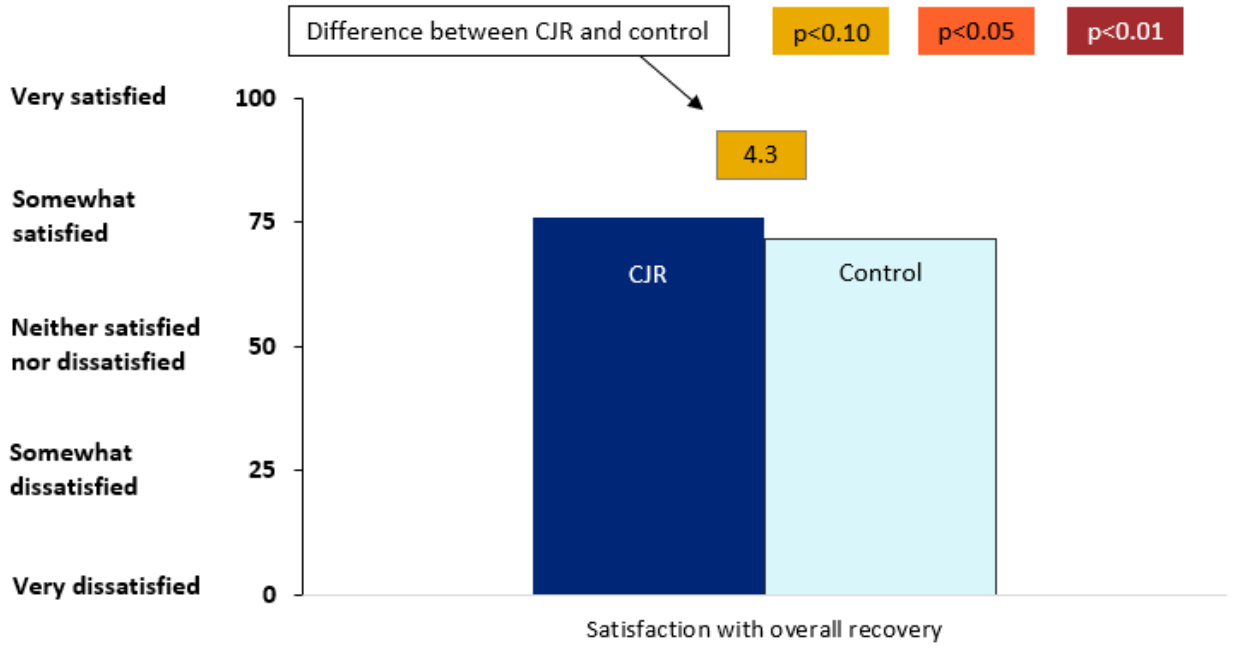
Satisfaction

More than 50% of respondents with hip fractures in the CJR and control groups were very satisfied with recovery since leaving the hospital, while less than 15% reported any level of dissatisfaction. Average satisfaction with recovery was just over 75 points on a 100-point scale

[49] All parenthetical results in this section refer to differences between CJR and control respondents for a given measure, not measure levels for either group of respondents.

for CJR respondents and 72 points for control respondents, both of which are noticeably lower than the corresponding average satisfaction for respondents with elective LERJs (80 and 78, respectively; **Appendix E: Patient Survey Results, Exhibits E-6 and E-7).** The difference between CJR and control respondents with hip fractures was 4.3 points with CJR hip fracture respondents reporting higher satisfaction with recovery than did control hip fracture respondents (p=0.08) (**Exhibit III-16**). This is roughly equivalent to 1 additional CJR respondent out of every 25 indicating one category higher satisfaction than a corresponding control respondent. For example, 1 additional CJR respondent may have indicated "very satisfied," while the corresponding control respondent only indicated "somewhat satisfied."

Exhibit III-16: CJR Respondents With Hip Fractures Reported More Satisfaction With Overall Recovery Than Control Respondents With Hip Fractures



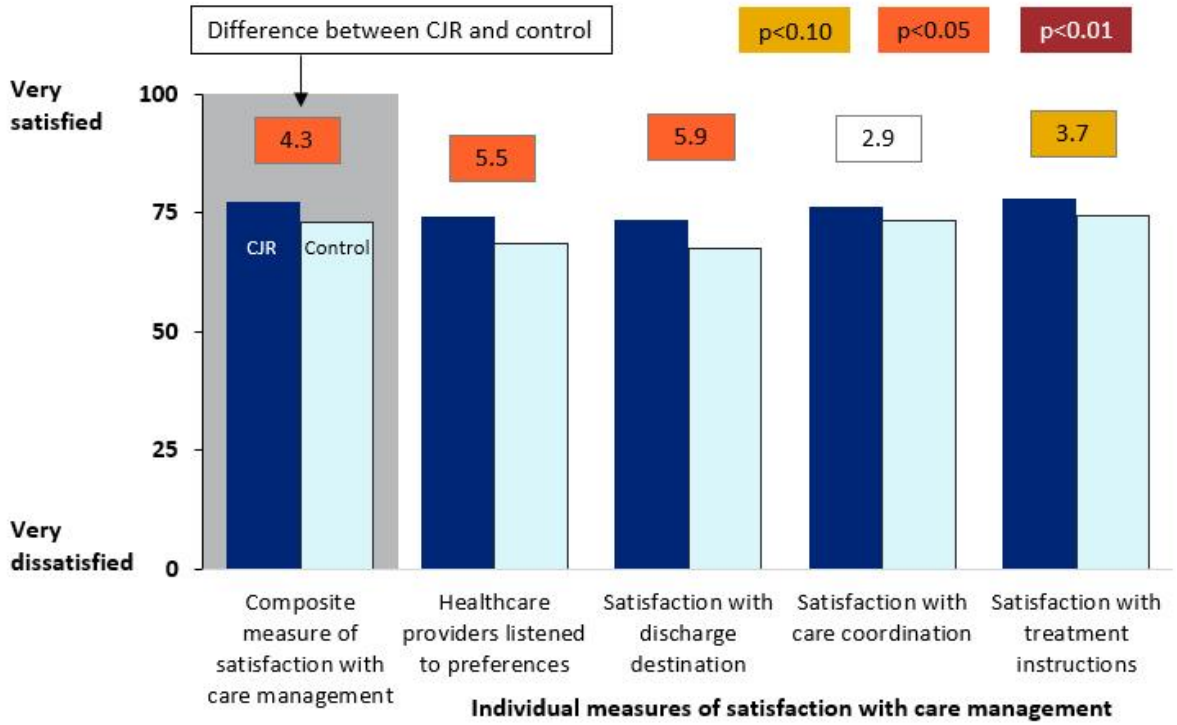
Source: CJR evaluation team analysis of patient survey data for episodes with inpatient discharge or outpatient procedure in July – October 2022.

Notes: The vertical bars depict the risk-adjusted average for CJR and control respondents. The difference between CJR and control is reported at the top of the figure in boxes. Measures of satisfaction consisted of a Likert scale with five levels and were normalized so that the lowest response category ("very dissatisfied") yielded a score of 0 and the highest response category ("very satisfied") yielded a score of 100. The estimates in this exhibit are the results of a cross-sectional regression model, weighted for sampling and nonresponse. See **Appendix B: Data and Methods, Exhibit B-15** for the list of covariates included in the regression and weighting procedures. Estimates that are significant at the 1%, 5%, or 10% significance level are indicated by red, orange, or yellow shading, respectively. CJR = Comprehensive Care for Joint Replacement.

Average satisfaction with care management among hip fracture respondents was around 75 points on a 100-point scale among CJR respondents, which was 4–5 points higher than among control respondents (**Exhibit III-17**). CJR respondents reported significantly higher satisfaction according to a composite measure across all measures of satisfaction with care management (4.3 points; p=0.03) and reported significantly higher satisfaction for three of the four individual measures. Specifically, they reported significantly higher satisfaction with the extent to which providers listened to their preferences about medical treatment (5.5 points; p=0.01), with discharge destination (5.9 points; p=0.01), and with treatment instructions (3.7 points; p=0.06). These results

are roughly equivalent to 4 to 5 additional CJR respondents out of 100 reporting that they were "somewhat" or "very" satisfied instead of "somewhat" or "very" dissatisfied for each measure.

Exhibit III-17: CJR Respondents With Hip Fractures Reported More Satisfaction With Care Management Than Control Respondents With Hip Fractures



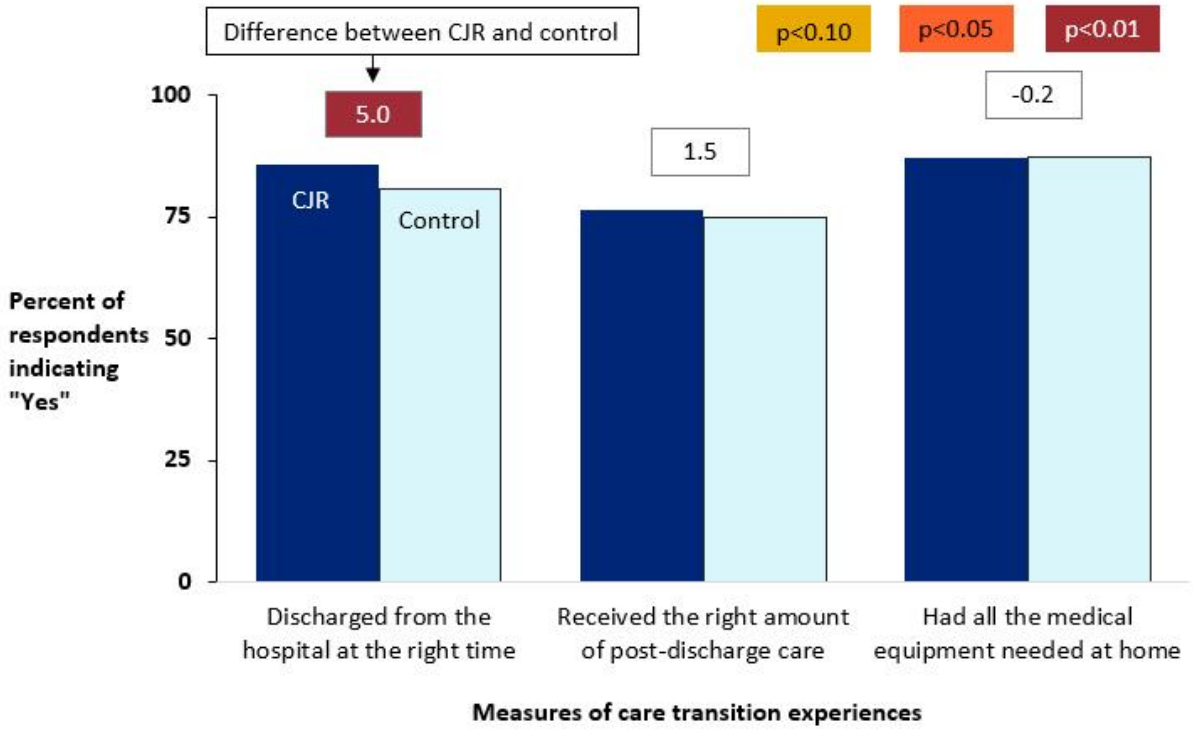
Source: CJR evaluation team analysis of patient survey data for episodes with inpatient discharge or outpatient procedure in July – October 2022.

Notes: The vertical bars depict the risk-adjusted average for CJR and control respondents. The difference between CJR and control is reported at the top of the figure in boxes. Measures of satisfaction consisted of a Likert scale with five levels and were normalized so that the lowest response category ("very dissatisfied") yielded a score of 0 and the highest response category ("very satisfied") yielded a score of 100. The estimates in this exhibit are the results of a cross-sectional regression model, weighted for sampling and nonresponse. See **Appendix B: Data and Methods, Exhibit B-15** for the list of covariates included in the regression and weighting procedures. Estimates that are significant at the 1%, 5%, or 10% significance level are indicated by red, orange, or yellow shading, respectively. CJR = Comprehensive Care for Joint Replacement.

Experience With Care Transitions

Most respondents with hip fractures in both the CJR and control groups reported having a positive experience with care transitions across all three measures (**Exhibit III-18**). CJR respondents were more likely than control respondents to agree they were discharged from the hospital at the right time (5.0 percentage points; $p<0.01$).

Exhibit III-18: CJR Respondents With Hip Fractures Were More Likely to Agree They Were Discharged From the Hospital at the Right Time



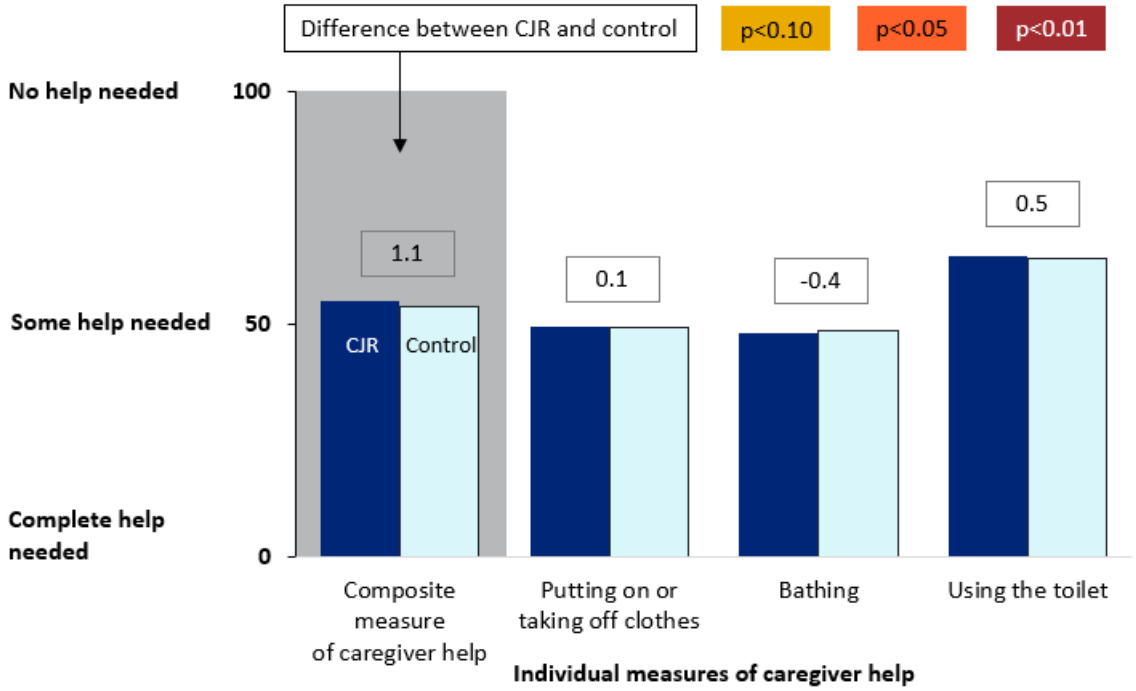
Source: CJR evaluation team analysis of patient survey data for episodes with inpatient discharge or outpatient procedure in July – October 2022.

Notes: The vertical bars depict the risk-adjusted average for CJR and control respondents. The difference between CJR and control is reported at the top of the figure in boxes. The estimates in this exhibit are the results of a cross-sectional regression model, weighted for sampling and nonresponse. See **Appendix B: Data and Methods, Exhibit B-15** for the list of covariates included in the regression and weighting procedures. Estimates that are significant at the 1%, 5%, or 10% significance level are indicated by red, orange, or yellow shading, respectively. CJR = Comprehensive Care for Joint Replacement.

Caregiver Help

About 96% of both CJR and control respondents with hip fractures received help from a caregiver after returning home. Among those who received help from a caregiver, most needed at least some help with bathing and dressing, while about half needed help using the toilet. Roughly one in four respondents reported "complete" help needed with at least one measure. We observed no significant differences between CJR and control respondents with hip fractures in the amount of help needed from caregivers (**Exhibit III-19**).

Exhibit III-19: CJR and Control Respondents With Hip Fractures Needed Similar Amounts of Caregiver Help at Home



Source: CJR evaluation team analysis of patient survey data for episodes with inpatient discharge or outpatient procedure in July – October 2022.

Notes: The vertical bars depict the risk-adjusted average for CJR and control respondents. The difference between CJR and control is reported at the top of the figure in boxes. Measures consisted of a Likert scale with five levels and were normalized so that the lowest response category ("complete help needed") yielded a score of 0 and the highest response category ("no help needed") yielded a score of 100. The estimates in this exhibit are the results of a cross-sectional regression model, weighted for sampling and nonresponse. See **Appendix B: Data and Methods, Exhibit B-15** for the list of covariates included in the regression and weighting procedures. Estimates that are significant at the 1%, 5%, or 10% significance level are indicated by red, orange, or yellow shading, respectively. CJR = Comprehensive Care for Joint Replacement.

d. Sensitivity Analysis

We conducted additional analyses to understand the extent to which overlap from the Bundled Payments for Care Improvement (BPCI) Advanced model may have influenced our survey results. As discussed in **Appendix E: Patient Survey Results**, 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 19.5% of control respondents across all LEJR episodes were treated by BPCI Advanced providers participating in LEJR (12.6% of control respondents with hip fractures were treated by BPCI Advanced participants). This overlap may bias our results if outcomes from hospitals in BPCI Advanced differ from what they would have been in the absence of BPCI Advanced. To explore this concern, we replicated each of the survey analyses, for all LEJRs, inpatient elective LEJRs, outpatient elective LEJRs, and hip fracture LEJRs, excluding any control respondents treated by a BPCI Advanced participant. We report these results in **Appendix E: Patient Survey Results**, **Exhibits E-10, E-11, and E-12**. Results were substantively the same, indicating that potential overlap from BPCI Advanced is not materially influencing our results. Notably, overlap with BPCI Advanced does not account for differences in functional status among CJR and control respondents with hip fractures.

IV. Medicare Program Savings

Acronyms	
ADI	Area Deprivation Index
AR	Annual Report
CJR	Comprehensive Care for Joint Replacement
CMS	Centers for Medicare & Medicaid Services
DiD	difference-in-differences
DSH	Disproportionate Share Hospital
LEJR	lower extremity joint replacement
MSA	metropolitan statistical area
M	million
N/A	not applicable
NPRA	net payment reconciliation amount
PY	performance year
SNH	safety-net hospital

Medicare achieves savings under the Comprehensive Care for Joint Replacement (CJR) Model when the relative reductions in episode payments at CJR hospitals exceed the net reconciliation payments made from the Centers for Medicare & Medicaid Services (CMS) to CJR hospitals. Over the past five performance years (PYs), the CJR Model resulted in overall estimated losses to Medicare, driven primarily by voluntary opt-in hospitals—hospitals that chose to continue as CJR participants after PY 2. Mandatory CJR hospitals—hospitals required to continue as CJR participants after PY 2—generated overall estimated savings in the first 4 PYs, but increased reconciliation payments in PY 5 eliminated the savings. The CJR Model evaluation's Fifth Annual Report (AR 5) presents detailed results for previous PYs.[50]

CMS made substantial changes to the CJR Model in PY 6 that may have improved the model's ability to generate estimated savings:

1. The model required all voluntary CJR hospitals, those participating in voluntary metropolitan statistical areas (MSAs), and low-volume and rural hospitals participating in mandatory MSAs to stop participating. As a result, our analyses in this chapter cover only mandatory CJR participants in PY 6.
2. CMS added episode-based payments for outpatient lower extremity joint replacements (LEJRs) to the model.
3. CMS adjusted hospital target prices based on age, dual-eligibility status, count of Hierarchical Condition Categories, and market trends, in addition to using a discount factor that adjusts for composite quality score as in prior years. The new risk adjustments and inclusion of outpatient LEJR procedures may have led to more accurate target prices for mandatory CJR hospitals.

In addition to these changes, PY 6 used 2019, the fourth PY of the CJR Model, as the historical period to determine target prices for hospitals. Thus, target prices for PY 6 incorporated payment reductions achieved by hospitals in 2019, and this may have also contributed to more accurate target prices.

This chapter presents estimated Medicare savings for PY 6, which includes episodes that ended between October 1, 2021, and December 31, 2022. In addition to Medicare savings, reconciliation payments to and repayments from hospitals indicates how hospitals fared, especially in the context of changes implemented in PY 6. We analyzed the distribution of the net reconciliation payment amounts (NPRA) and how the NPRA distribution differs from prior PYs.

[50] Centers for Medicare & Medicaid Services. (2023). *Comprehensive Care for Joint Replacement Model - Fifth Annual Report*. https://www.cms.gov/priorities/innovation/data-and-reports/2023/cjr-py5-annual-report

A. After Accounting for Reconciliation Payments, Did the CJR Model Result in Medicare Savings?

1. Summary of Findings

The callout box presents a summary of our findings. Results with an asterisk (*) are key findings.

- (*) During PY 6 (2021–2022), the CJR Model resulted in statistically significant estimated savings of $54.2 million (M). The savings may have ranged from $2.0 M to $106.4 M based on the 90% confidence interval.
- (*) The CJR Model saved an estimated $1,017 per episode, which may have ranged from $38 to $1,996 based on the 90% confidence interval.
- The distribution of reconciliation and repayments across hospitals in PY 6 remains highly unequal. The top 10% of hospitals received more than $18 M in reconciliation, while the bottom 10% of hospitals repaid more than $19 M.
- Between PY 1–5 and PY 6, many hospitals moved within the NPRA distribution. A substantial portion of hospitals that received large amounts of NPRA in PY 1–5 fell to the bottom of the distribution and repaid large amounts to CMS in PY 6.

2. Methods

To examine the impact of the model on Medicare savings and changes in reconciliation payments, we conducted three analyses:

- Estimated Medicare program savings in PY 6
- Examined the distribution of reconciliation payments across CJR participants
- Assessed changes in the distribution of reconciliation payments between PY 6 and prior PYs

a. Medicare Program Savings Estimation in PY 6

We calculated Medicare savings on both a total and a per-episode basis by subtracting net reconciliation payments to CJR participant hospitals from the change in nonstandardized paid amounts due to the CJR Model.

Medicare savings = change in nonstandardized paid amounts – net reconciliation payments

Change in Nonstandardized Paid Amounts

Estimates from a difference-in-differences (DiD) model of per-episode standardized paid amounts generated the change in nonstandardized paid amounts. We multiplied the DiD estimates by –1 and converted them to nonstandardized paid amounts using a ratio of nonstandardized to standardized Medicare paid amounts from PY 6 CJR episodes.[51] Multiplying the per-episode estimate by the total

[51] **Appendix B: Data and Methods** provides more details on the conversion from standardized to nonstandardized paid amounts.

number of CJR episodes in the regression sample generated the total change in nonstandardized paid amounts. We calculated the ranges that we report as confidence intervals in a similar fashion from the 90% confidence interval of the DiD model of per-episode standardized paid amounts.

In Medicare savings calculations, we used nonstandardized paid amounts instead of the standardized allowed amounts used in the average episode and service-level payment impact analyses. Nonstandardized paid amounts represent the actual payments from Medicare to providers, incorporating geographic and other payment adjustments and excluding beneficiary cost-sharing. Using nonstandardized paid amounts aligns with the calculation of reconciliation payments. As a result, the change in paid amounts reported in this chapter differs from the change in allowed amounts reported in **Chapter III: Impact of the Model**. In general, the change in paid amounts is smaller than the change in allowed amounts because it does not include the change in beneficiary cost-sharing.

Reconciliation Payments

Reconciliation payments represent the net repayments from CJR participants to Medicare, either positive or negative. The CJR Final Rule refers to them by the term *net payment reconciliation amounts*, or NPRA.[52] The CMS CJR payment contractor provided these data. We calculated reconciliation payments per episode by dividing total reconciliation payments by the total number of CJR episodes.

Net reconciliation payments equal the sum of payments made to CJR participant hospitals by Medicare for meeting cost and quality targets, less summed repayments from CJR participant hospitals to Medicare for failing to meet cost and quality targets.

b. Distribution of PY 6 Reconciliation Payments Across CJR Participants

We analyzed the distribution of NPRA across hospitals by ordering hospitals by their PY 6 NPRA amounts and then grouping the ordered hospitals into deciles. Hospitals in lower-valued deciles received more in reconciliation or repaid less to CMS than hospitals in higher-valued deciles. For each decile, we calculated the total reconciliation amount for all the hospitals in the decile as well as the average net reconciliation per hospital. We performed a similar analysis for total PY 1–5 NPRA in AR 5.[53] Additionally, we conducted exploratory analyses to study the characteristics of hospitals that received the largest reconciliation payments or had the largest repayments using descriptive statistics (percentages, counts, averages). For this exploratory study, we looked at the number of PY 6 episodes, the proportion of hospitals affiliated with a health system, the percentage of hospitals identified as safety-net hospitals (SNHs) under various definitions, the composition of quality ratings, and the quality discount rate.

[52] Medicare Program; Comprehensive Care for Joint Replacement Payment Model for Acute Care Hospitals Furnishing Lower Extremity Joint Replacement Services, 80 Fed. Reg. 73273 (November 24, 2015) (codified at 42 CFR 510).

[53] Centers for Medicare & Medicaid Services. (2023). *Comprehensive Care for Joint Replacement Model - Fifth Annual Report.* https://www.cms.gov/priorities/innovation/data-and-reports/2023/cjr-py5-annual-report

c. Changes in Distribution of Reconciliation Payments Between PY 6 and Prior PYs

Given the substantial changes to model rules described above, we analyzed the movement of hospitals within the NPRA distribution from PY 1–5 to PY 6. For this analysis, we calculated the count and percentage of hospitals that had the same or different NPRA decile groups in PY 1–5 and PY 6. Additionally, for the subset of hospitals that moved from the top of the NPRA distribution in PY 1–5 (received reconciliation) to the bottom of the distribution in PY 6 (owed repayments), we conducted exploratory analyses to describe the hospital characteristics using descriptive statistics.

3. Results

a. Medicare Program Savings Estimation in PY 6

In PY 6, the CJR Model resulted in Medicare savings of $54.2 M for all episodes in PY 6, with a range based on the 90% confidence interval of $2.0 M to $106.4 M (**Exhibit IV-1**). On average, the model led to savings of $1,017 per episode, with a range based on the 90% confidence interval of $38 to $1,996. Medicare payments decreased by $939 per episode, with a 90% confidence interval ranging from an increase of $40 to a reduction of $1,918. The average net payment reconciliation per episode was –$78, indicating net repayments from hospitals to CMS. PY 6 represents the first PY of the CJR Model that resulted in average net repayments from mandatory hospitals. In previous PYs, the average mandatory hospital received reconciliation payments from CMS. The estimated reduction in payments of $939 per episode is larger than the recent reductions in PY 4 and PY 5, almost reaching the estimated reduction for PY 3.

Exhibit IV-1: During PY 6, Mandatory CJR Hospitals Are Estimated to Have Generated $54.2 M in Medicare Savings

Savings component	Value	90% confidence interval
Reduction in nonstandardized paid amounts per episode	$939	-$40 to $1,918
Reconciliation payments per episode	-$78	N/A
Medicare savings per episode	$1,017	$38 to $1,996
Number of PY 6 episodes	53,328	N/A
Aggregate Medicare savings	$54,228,811	$2,013,099 to $106,444,549

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 ending between October 1, 2021, and December 31, 2022 (PY 6 intervention), as well as CJR payment contractor data for CJR participant hospitals in PY 6.

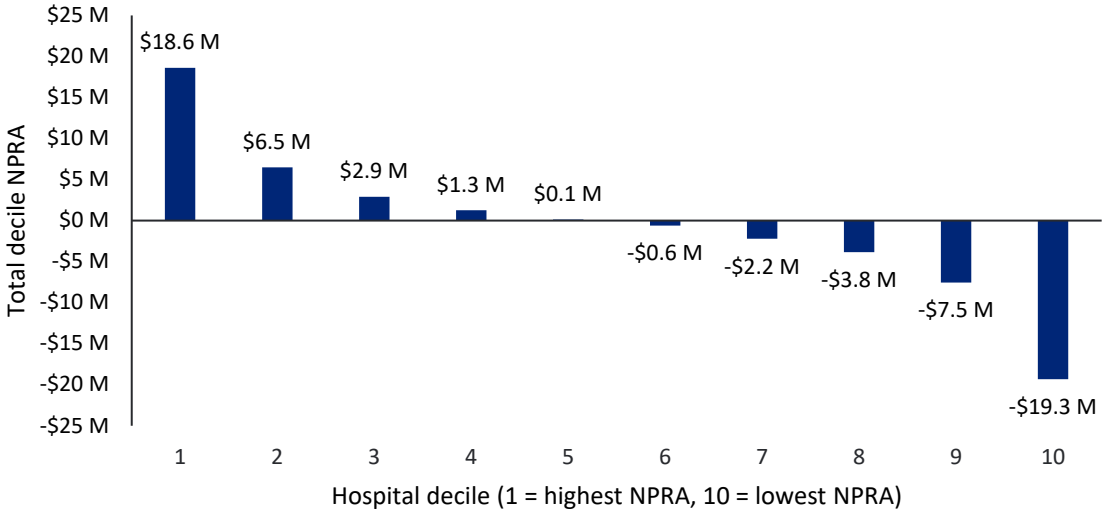
Notes: Reductions in nonstandardized paid amounts are based on a weighted average of PY estimates from a DiD model of per-episode standardized paid amounts that have been multiplied by –1 and converted to nonstandardized amounts. We do not report confidence intervals for reconciliation payments per episode and number of PY 6 episodes because these were not estimated but observed with certainty. We calculated aggregate Medicare savings by multiplying Medicare savings per episode by the number of PY 6 episodes, and Medicare savings per episode equals the estimated reduction in nonstandardized paid amounts per episode minus the average reconciliation payments per episode. Because reconciliation payments per episode depend on realized payment reductions, the asterisked intervals are not truly confidence intervals but ranges based on the confidence interval. CJR = Comprehensive Care for Joint Replacement; DiD = difference-in-differences; M = million; N/A = not applicable; PY = performance year.

b. Distribution of Reconciliation Payments in PY 6 Across CJR Participants

We found large variations in NPRA across hospitals when grouping hospitals by decile (**Exhibits IV-2 and V-3**). While the top 10% of hospitals received large reconciliation payments, a total of $18.6 M in PY 6, the bottom 10% of hospitals had a larger amount in repayments, totaling

$19.3 M in repayments. Hospitals that made up the middle 80% of the distribution averaged $13,500 in repayments to CMS per hospital. The general distribution is like that of PY 1–5, but it shifts toward repayments to CMS due to the substantially lower average NPRA in PY 6.

Exhibit IV-2: In PY 6, the Top 10% of CJR Hospitals Had Large Gains in NPRA, While the Bottom 10% Had Large Losses



Source: CJR evaluation team analysis of CJR payment contractor data for CJR participant hospitals in PY 6.

Notes: Distribution of total NPRA across hospital deciles. Decile 1 contains the top 10% of PY 6 participant hospitals in terms of NPRA receipt, while decile 10 contains the bottom 10% of PY 6 participant hospitals in terms of NPRA receipt. Data labels indicate the cumulative reconciliation received by hospitals in a given decile. CJR = Comprehensive Care for Joint Replacement; M = million; NPRA = net payment reconciliation amounts; PY = performance year.

Exhibit IV-3: High-NPRA Hospitals in PY 6 Were More Likely to be Affiliated With a Hospital System, Were More Likely to Have More LEJR Patients From High-ADI Neighborhoods, and Had a Lower Discount Rate Adjustment for Composite Quality Than Low-NPRA Hospitals

Outcome	Hospitals in the bottom 30%		Hospitals in the top 30%	
Net reconciliation in PY 6	Repayment of $320,069		Reconciliation of $291,705	
Average number of PY 6 episodes	240		226	
Percentage affiliated with a health system	77%		85%	
Number defined as safety-net hospitals	24 DSH, 22 dual%, 16 ADI		11 DSH, 7 dual%, 28 ADI	
Composite quality ratings	Excellent	9	Excellent	19
	Good	54	Good	68
	Acceptable	19	Acceptable	9
	Below acceptable	14	Below acceptable	0
Average composite quality discount	1.9%		1.3%	

Source: CJR evaluation team analysis of Medicare claims and enrollment data for episodes ending between October 1, 2021, and December 31, 2022 (PY 6). CJR payment contractor data for CJR participant hospitals in PY 6, and the Fiscal Year 2022 Inpatient Prospective Payment System Final Rule and Correction Notice.

Notes: The top and bottom 30% of hospitals in the PY 6 NPRA distribution each comprise 96 hospitals. We used three potential definitions of safety-net hospitals: 85th percentile or above nationally in Disproportionate Share Percentage (DSH), at least 40% of all patients having dual eligibility for Medicare and Medicaid (dual%), or at least 30% of LEJR recipients having an ADI score greater than 70 (ADI). ADI = Area Deprivation Index; CJR = Comprehensive Care for Joint Replacement; DSH = Disproportionate Share Hospital; LEJR = lower extremity joint replacement; NPRA = net payment reconciliation amounts; PY = performance year.

Hospitals in the top 30% of the PY 6 NPRA distribution received an average of $291,705 in reconciliation payments from CMS. Hospitals in the bottom 30% paid an average of $320,069 back to CMS. Both hospital groups were similar in terms of PY 6 episode volume. The top 30% of hospitals were more likely to be affiliated with a health system, more likely to be defined as an SNH based on the Area Deprivation Index, and less likely to be defined as an SNH based on Disproportionate Share Percentage or proportion of patients with dual eligibility.

c. *Changes in Distribution of Reconciliation Payments Between PY 6 and Prior PYs*

We also looked at how CJR hospitals fared under the new target pricing in PY 6 by comparing their relative NPRA before and during PY 6. In relative terms, there was notable movement within the NPRA distribution between PY 1–5 and PY 6 (**Exhibit IV-4**). Of note, a substantial portion of hospitals in the top 30% of the PY 1–5 NPRA distribution fell to the bottom 30% of the PY 6 NPRA distribution, while many hospitals in the middle of the PY 1–5 distribution rose, in relative terms, to higher-NPRA deciles in the PY 6 distribution. Hospitals from the bottom of the PY 1–5 NPRA distribution tended to stay in the same decile or rise only slightly in the distribution.

Exhibit IV-4: High-NPRA Hospitals in PY 6 Tended to Be Relatively Lower in the NPRA Distribution in PY 1–5, and a Substantial Portion of Low-NPRA Hospitals in PY 6 Were High-NPRA Hospitals in PY 1–5

		High NPRA ← PY 6 decile → Low NPRA									
		1	2	3	4	5	6	7	8	9	10
High NPRA	1	44%	6%	6%	0%	3%	3%	6%	3%	9%	19%
	2	25%	13%	6%	3%	6%	0%	6%	25%	9%	6%
	3	28%	13%	6%	13%	6%	3%	3%	6%	9%	13%
	4	3%	34%	16%	19%	13%	3%	3%	0%	3%	6%
PY 1–5 decile	5	0%	16%	25%	13%	13%	13%	6%	9%	3%	3%
	6	0%	6%	13%	25%	19%	9%	9%	9%	6%	3%
	7	0%	0%	6%	19%	13%	13%	25%	13%	6%	6%
	8	0%	9%	6%	6%	13%	34%	25%	3%	0%	3%
Low NPRA	9	0%	0%	13%	0%	13%	16%	13%	13%	19%	16%
	10	0%	3%	3%	3%	3%	6%	3%	19%	34%	25%

Row/Column Color Legend

White: Hospitals received reconciliation payments

Tan: Hospitals repaid money to CMS

Cell Color Legend

Green: Hospitals were in a lower decile in PY 6 relative to PY 1–5 ("upward mobility")

Yellow: Hospitals were in the same decile in PY 6 relative to PY 1–5

Red: Hospitals were in a higher decile in PY 6 relative to PY 1–5 ("downward mobility")

Source: CJR evaluation team analysis of CJR payment contractor data for CJR participant hospitals in PY 1–6.

Notes: Matrix shows a hospital's position in the NPRA distribution in PY 6 and in PY 1–5. Rows indicate the PY 1–5 decile, and columns indicate the PY 6 decile. Values in each cell indicate what percentage of all PY 6 hospitals were in the row's PY 1–5 decile and the column's PY 6 decile. CJR = Comprehensive Care for Joint Replacement; NPRA = net payment reconciliation amounts; PY = performance year.

Among the hospitals in the top 30% of the PY 1–5 NPRA distribution, 49% remained in the top 30% of the PY 6 NPRA distribution. However, 33% of the hospitals in the top 30% of the distribution in PY 1–5 fell to the bottom 30% of the PY 6 distribution.

The hospitals that fell from the top 30% of the PY 1–5 NPRA distribution to the bottom 30% of the PY 6 NPRA distribution tended to have fewer beds and higher PY 6 episode volume than the average hospital in the top 30% of the PY 1–5 NPRA distribution (**Exhibit IV-5**).

Exhibit IV-5: Hospitals From the Top 30% in PY 1–5 That Fell to the Bottom 30% in PY 6 Had Lower Average Bed Counts and Higher Average PY 6 Episode Volume Than the Average Hospital in the Top 30% in PY 1–5

Outcome	Hospitals in the top 30% in PY 1–5 that fell to the bottom 30% in PY 6	All Hospitals in the top 30% in PY 1–5
Number of hospitals	32	96
Average bed count	251	294
Average PY 6 episodes	451	376
Average PY 6 NPRA	-$371,989	$92,812

Source: CJR evaluation team analysis of CJR payment contractor data for CJR participant hospitals in PY 1–6.

Notes: CJR = Comprehensive Care for Joint Replacement; NPRA = net payment reconciliation amounts; PY = performance year.

Among hospitals in the bottom 30% of the PY 1–5 NPRA distribution, 44% remained in the bottom 30% of the PY 6 NPRA distribution. A small proportion (11%) of hospitals shifted to the top 30%. The remaining 45% were in the middle 40% of the PY 6 NPRA distribution.

V. Health Equity Impact of the CJR Model

In its 2022–2032 Framework for Health Equity, the Centers for Medicare & Medicaid Services (CMS) defines *health equity* as "the attainment of the highest level of health for all people, where everyone has a fair and just opportunity to attain their optimal health regardless of race, ethnicity, disability, sexual orientation, gender identity, socioeconomic status, geography, preferred language, or other factors that affect access to care and health outcomes."[54] One of the priorities includes explicitly measuring the impact of policies on health equity to inform the development of sustainable solutions that close gaps in health and in health care access, quality, and outcomes. Additionally, CMS set a goal to evaluate policies to determine how CMS can support safety-net providers, including acute care hospitals, as part of their agenda for advancing health equity.[55, 56]

Acronyms	
ADI	Area Deprivation Index
CJR	Comprehensive Care for Joint Replacement
DiD	difference-in-differences
DDD	difference-in-difference-in-differences
ED	emergency department
ESRD	end-stage renal disease
FFS	fee-for-service
HCC	Hierarchical Condition Category
HH	home health
IRF	inpatient rehabilitation facility
LCI	lower confidence interval
LEJR	lower extremity joint replacement
MSA	metropolitan statistical area
p	p-value
PAC	post-acute care
pp	percentage point
PY	performance year
SDOH	social determinants of health
SNF	skilled nursing facility
SNH	safety-net hospital
UCI	upper confidence interval

In support of this goal, we conducted analyses to examine the impact of the CJR Model on underserved populations. CMS defines *underserved populations* as "populations sharing a particular characteristic, including geographic communities, that have been systematically denied a full opportunity to participate in aspects of economic, social, and civic life."

Implemented in 2016, the CJR Model did not explicitly include principles of health equity in the model design. Nonetheless, changes in payments to providers and incentives to improve quality can have different impacts on diverse groups of beneficiaries. To study the impact of the CJR Model on underserved populations, we examined (1) differences, or "gaps," in health outcomes experienced by people disadvantaged or underserved at baseline and (2) whether CJR narrowed or widened these gaps. We conducted analyses for four underserved populations[57]:

1. Black or African American patients,
2. Patients who are dually eligible for Medicare and Medicaid,
3. Patients who are both Black or African American and dually eligible, and

[54] Centers for Medicare & Medicaid Services. (2022). *CMS Framework for Health Equity 2022–2032*. https://www.cms.gov/files/document/cms-framework-health-equity.pdf

[55] Centers for Medicare & Medicaid Services. (n.d.). *Health equity*. https://www.cms.gov/priorities/innovation/key-concepts/health-equity

[56] Centers for Medicare & Medicaid Services. (2023). CMS *strategic plan: health equity*. https://www.cms.gov/files/document/health-equity-fact-sheet.pdf

[57] The analyses of the changes in race and ethnicity rely on the Research Triangle Institute (RTI) race codes from the Master Beneficiary Summary File. The RTI race code is created based on beneficiaries' self-reporting to Medicare and the Social Security Administration and RTI's race imputation algorithm utilizing beneficiary names and geographical locations. Additional information about the RTI race code can be found at https://resdac.org/cms-data/variables/research-triangle-institute-rti-race-code.

4. Hispanic patients.

This approach benchmarks the outcomes for an underserved population to the outcomes for a reference population. In race and ethnicity analyses, researchers typically use beneficiaries who are Non-Hispanic White as the reference population. Underserved populations may have systematically different health conditions and health care needs than reference populations. Consequently, the optimal level of a health outcome may differ between the underserved population and the reference population; for example, beneficiaries who are dually eligible might have a higher level of appropriate emergency department (ED) visits than beneficiaries who are not dually eligible. It can be difficult to attach normative inferences to the gaps, or the changes in gaps, unless we know the "right" level of an outcome for a population.

We looked at the proportion of underserved patients in the model; model impacts on LEJR rates; and model impacts on various outcomes, including cost, quality, utilization, and patient-reported perception of functional status and care experiences (as presented in **Chapter III: Impact of the Model**).

We focus on elective LEJRs in our analysis of underserved populations because elective procedures offer greater opportunities for discretion and alternative care pathways. In contrast, for hip fractures a hip replacement is the typical clinical response, and the acute nature of a fracture gives hospitals limited discretion in terms of pre-surgery patient optimization or scheduling. Therefore, to examine the impact of the CJR Model on LEJR rates and other claims-based outcomes, we considered only elective LEJRs.

We explored how safety-net hospitals (SNHs) in our population (CJR participants) approached care for underserved populations and asked how CJR affected their strategies and success. We interviewed six CJR participants identified as a SNH to obtain broad perspectives on the relationship between the CJR Model and the hospitals' care strategies.

Each section in this chapter has three subsections: summary of findings, methods, and results. The methods describe unique elements of our approach specific to the research question. **Appendix B: Data and Methods** provides details on all the methods used in this chapter.

A. How Does the Proportion of Underserved Patients Compare for the CJR and Control Population?

In this section, we present findings on the proportion of patients from underserved populations in the CJR population compared with the control population in Performance Year 6 (PY 6). We examined how the size of the underserved population varied among Medicare fee-for-service (FFS) beneficiaries as well as for LEJR episodes across the CJR and control groups. To remain consistent with the rest of this chapter, we focused on the elective LEJR population.[58]

[58] A detailed table summarizing all four population groups can be found in **Appendix B: Data and Methods**.

1. Summary of Findings

- For two of the populations, beneficiaries who are Hispanic and beneficiaries who were dually eligible, we found that the proportion of underserved beneficiaries differs for CJR and control populations at the metropolitan statistical area (MSA) level. We did not observe this difference for the Black or African American or the Black or African American and dually eligible populations.

2. Methods

We examined the proportion of underserved populations in CJR and control for the four population groups defined in **Exhibit V-1**. We defined these groups based on Medicare FFS beneficiary residence or receipt of an LEJR in PY 6. Specifically, for each identified population group, we calculated the proportion of beneficiaries who are (1) Black or African American, (2) Hispanic, (3) dually eligible, and (4) both Black or African American and dually eligible.

We also examined the changes in proportion of underserved population at CJR and control hospitals over time. A detailed discussion of this analysis is in **Appendix F: Claims-Based Health Equity Analyses, Section B**.[59]

Exhibit V-1: Proportion of Underserved Beneficiaries and Episodes

Population definition criteria	Population definitions by CJR and control groups		Proportion calculation
	CJR	Control	
Beneficiary residing at CJR/control MSA in a year[a] **(01/2022–12/2022)**	• ***CJR MSA beneficiaries:*** Medicare FFS beneficiaries *residing* in a CJR MSA • ***CJR MSA beneficiaries with LEJR:*** CJR MSA beneficiaries who had at least one LEJR • ***CJR MSA beneficiaries with elective LEJR:*** CJR MSA beneficiaries who had at least one elective LEJR	• ***Control MSA beneficiaries:*** Medicare FFS beneficiaries residing in a control MSA • ***Control MSA beneficiaries with LEJR:*** control MSA beneficiaries who had at least one LEJR • ***Control MSA beneficiaries with elective LEJR:*** control MSA beneficiaries who had at least one elective LEJR	Denominator = number of FFS beneficiaries for the population Numerator = among denominator, number of FFS beneficiaries identified in each underserved population based on RTI race code or dual-eligibility indicator in Medicare enrollment data
LEJR episode[b] (episode ending between 10/2021 and 12/2022)	• ***CJR elective LEJR episodes:*** elective LEJR episodes at a participating CJR hospital	• ***Control elective LEJR episodes:*** elective LEJR episodes at a control hospital	Denominator = number of LEJR episodes Numerator = number of episodes related to patients identified in each underserved population based on RTI race code or dual-eligibility indicator in Medicare enrollment data

Source: CJR evaluation team analysis of Medicare claims and enrollment data for Medicare FFS beneficiaries enrolled in Medicare during 2022.

[59] Full details on all the methods used for the analyses in this section can be found in **Appendix B: Data and Methods**.

Notes: CJR = Comprehensive Care for Joint Replacement; ESRD = end-stage renal disease; FFS = fee-for-service; LEJR = lower extremity joint replacement; MSA = metropolitan statistical area; PY = performance year.

[a] Data source is a beneficiary/year-level analytic dataset capturing the list of all Medicare FFS beneficiaries, excluding beneficiaries eligible for Medicare based on ESRD status, living in mandatory CJR MSAs or their corresponding control MSAs in each calendar year between 2012 and 2022. The data include information on beneficiary sociodemographic characteristics as well as any LEJRs the beneficiary had in a year and was constructed from Medicare claims and enrollment data. For this data, the variables related to LEJR were constructed based on all LEJRs including those that were performed at an ambulatory surgical center. For this analysis, we only used data for the Calendar Year 2022 which aligns with most of PY 6.[60]

[b] Data source is an episode level analytic dataset capturing information on all episodes of care related to LEJRs performed at a participating CJR hospital or their corresponding control hospital between baseline and PY 6 (2012–2022). Note, some hospitals in the CJR MSAs are not participating in the model due to low-volume or rural designation starting from PY 6. Only episodes related to hospitals participating in CJR Model in PY 6 have been included. This analytic data has been used for the majority of the impact analyses included in this report. For this analysis, we only used data related to the PY 6 evaluation. Some LEJR episodes at CJR participating hospitals can be for Medicare FFS beneficiaries who are not residing at a CJR MSA while some beneficiaries who live in a CJR MSA could have had their LEJR episode at a non-participating CJR hospital. Additionally, certain LEJR episodes can be excluded from CJR Model.[61] Hence the counts generated from beneficiary/year data and LEJR episode data are slightly different.

3. Results

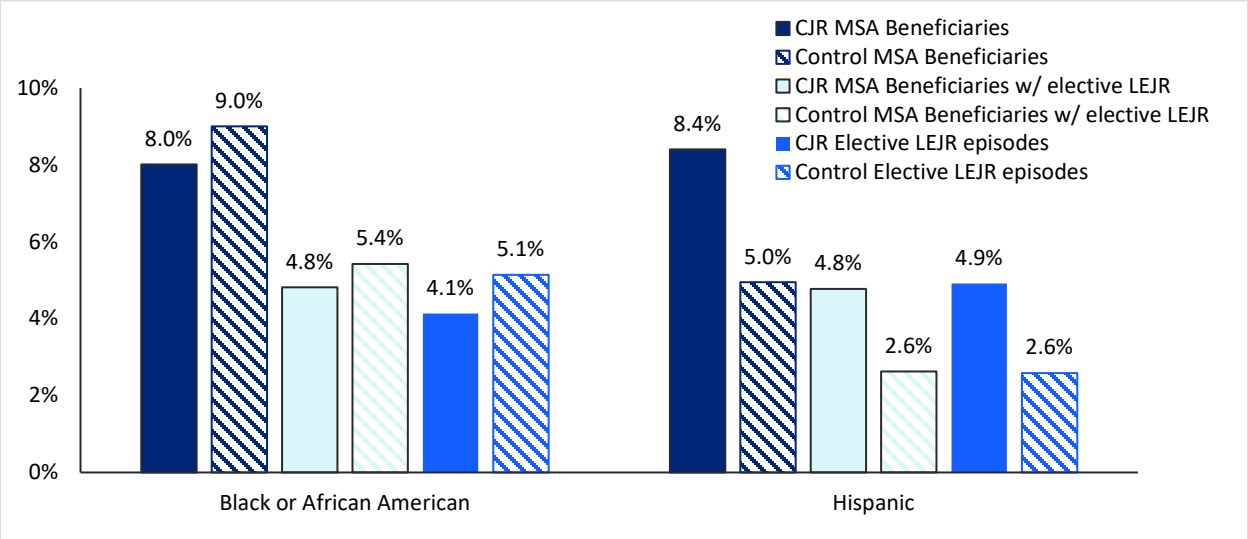
The proportion of underserved population varied based on the population definitions. Beneficiaries who are Black or African American made up 8.0% of the CJR MSA beneficiaries and 9.0% of the control MSA beneficiaries. In contrast, the proportion of beneficiaries who are Black or African American was 4.8% for CJR MSA beneficiaries with an elective LEJR and 5.4% for control MSA beneficiaries with an elective LEJR. The proportion of elective LEJR episodes for a patient who is Black or African American made up 4.1% of CJR elective LEJR episodes and 5.1% of control elective LEJR episodes (**Exhibit V-2**). Across the populations examined, we observed comparable proportions for the CJR and control groups.

Beneficiaries who are Hispanic made up 8.4% of FFS beneficiaries in CJR MSAs and 5.0% of beneficiaries for control MSAs. The proportion of beneficiaries who are Hispanic was lower for elective LEJR recipients in CJR MSAs (4.8%) and in control MSAs (2.6%). After applying model exclusion rules, elective LEJR episodes involving a patient who is Hispanic made up 4.9% of CJR episodes and 2.6% of control episodes (**Exhibit V-2**). Based on these results, we see that the difference in the proportion of episodes involving a patient who is Hispanic for the CJR and control populations were driven by differences in the beneficiary population in the two groups of MSAs.

[60] While PY 6 includes episodes ending between October 2021 and December 2022, the beneficiary-level analytic dataset captures information in a calendar year. Thus, there are episodes that started in the last 2 quarters of 2021 that are PY 6 episodes and not included in the beneficiary-level dataset for 2022, and there are LEJRs present in the beneficiary-level dataset that occurred in Q4 of 2022 that are not included in PY 6.

[61] LEJR episodes can be excluded from the CJR Model for various reasons, including alignment with certain ACOs and other Medicare coverage details, ending due to death within the 90-day post-discharge window or the patient being readmitted during an episode and discharged under MS-DRG 469, 470, 521, or 522 (in which case the first episode was cancelled and a new CJR episode began).

Exhibit V-2: Differences in the Proportion of Episodes Involving Patients Who Are Black or African American or Hispanic Between the CJR and Control Groups Were Comparable to Differences in the Beneficiary Population in CJR and Control MSAs



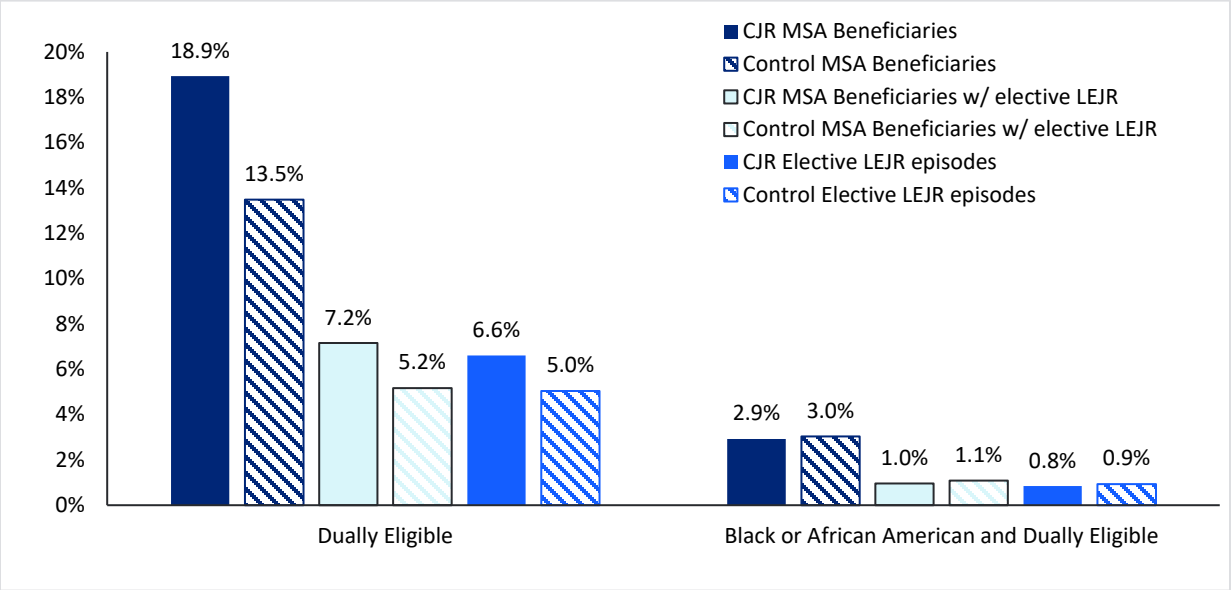
Source: CJR evaluation team analysis of Medicare claims and enrollment data for Medicare FFS beneficiaries enrolled in Medicare during 2022.

Notes: Proportions of FFS beneficiaries, FFS beneficiaries who received at least one LEJR, and elective CJR LEJR patients who are Black or African American or Hispanic. CJR = Comprehensive Care for Joint Replacement; FFS = fee-for-service; LEJR = lower extremity joint replacement; MSA = metropolitan statistical area.

Beneficiaries who were dually eligible made up 18.9% of the Medicare FFS beneficiary population in CJR MSAs and 13.5% in control MSAs. Beneficiaries who were dually eligible made up 7.2% of elective LEJR recipients in CJR MSAs and 5.2% in control MSAs. Applying model exclusion rules, elective LEJR episodes involving a patient who was dually eligible made up 6.6% of CJR episodes and 5.0% of control episodes (**Exhibit V-3**). We observed a similar pattern in the results for patients who are Hispanic. The differences between the proportion of episodes involving a patient who was dually eligible between CJR and control episode samples were driven by differences in the MSA populations between the CJR and control groups.

Beneficiaries who are both Black or African American and dually eligible made up 2.9% of the Medicare FFS beneficiary population in CJR MSAs and 3.0% of Medicare FFS beneficiaries in control MSAs, and they made up 1.0% of LEJR recipients in CJR MSAs and 1.1% of LEJR recipients in control MSAs. Applying model exclusion rules, elective inpatient LEJR episodes involving a patient who is both Black or African American and dually eligible made up 0.8% of CJR episodes and 0.9% of control episodes (**Exhibit V-3**). The proportions for the Black or African American and dually eligible population were comparable across the CJR and control groups.

Exhibit V-3: Beneficiaries Who Were Dually Eligible Were More Common in CJR MSAs Than Control MSAs, and the Share of Patients Who Are Both Black or African American and Dually Eligible Was Comparable Across Populations for the CJR and Control Groups



Source: CJR evaluation team analysis of Medicare claims and enrollment data for Medicare FFS beneficiaries enrolled in Medicare during 2022.

Notes: Proportions of FFS beneficiaries, FFS beneficiaries who received at least one LEJR, and elective CJR LEJR patients who dually eligible or are Black or African American and dually eligible. CJR = Comprehensive Care for Joint Replacement; FFS = fee-for-service; LEJR = lower extremity joint replacement; MSA = metropolitan statistical area.

Taken together, these results indicate that the largest drop in proportions across all underserved populations occurred because of the restriction applied from the Medicare FFS beneficiary population to the elective LEJR patient population. The difference between the proportion of underserved patients who received elective LEJRs and the proportion of underserved patients who had a CJR episode within CJR and control groups were generally within 1 percentage point of each other.

B. What Care Strategies Did CJR SNHs Employ?

One of CMS' strategies to advancing health equity is to identify ways to support safety-net providers, including acute care hospitals. For this evaluation, we: (1) examined the different definitions to identify a safety net hospital (SNH), (2) identified how many CJR participants can be classified as an SNH and, (3) interviewed six SNHs to understand their experience with the CJR Model and with providing care to patients with unmet nonmedical needs. In this section, we present findings from the analyses.

1. Summary of Findings

The callout box presents a summary of our findings. Results with an asterisk (*) are key findings.

- (*) Interviewees from SNHs expressed concerns about their financial prospects in the CJR Model, with particular concerns around the CJR target price being "too low" and not reflective of the high costs of care for their complex patient population.
- SNHs serve complex patient populations with high unmet medical and nonmedical needs. Care transformation for SNHs focused on social determinants of health (SDOH) rather than post-acute care (PAC) use.

2. Methods

a. Defining Safety-Net Hospitals

The term *safety-net hospital* does not have a single definition. In common use, the term refers to hospitals that provide services to individuals regardless of their insurance or ability to pay. In empirical studies, SNHs are most often identified using metrics like disproportionate share percentage measures of Medicaid caseload or the amount of uncompensated care that qualifies a hospital for additional Medicare payment. Less often, metrics related to dual eligibility of hospital patients or neighborhood deprivation measures such as the Area Deprivation Index (ADI) are used to define SNHs.

Given these varying definitions, we opted to identify SNHs using three approaches[62]:

- Disproportionate share percentage: A hospital is an SNH if it is in the top 15% of the national ranking of hospitals in terms of disproportionate share percentage.
- Dually eligible percentage: A hospital is an SNH if at least 40% of all hospital patients are enrolled in both Medicare and Medicaid.
- ADI: A hospital is an SNH if at least 30% of LEJR patients live in a 9-digit zip-code with ADI greater than 70.

Applied to the set of hospitals included in our analyses, these definitions identified the following SNHs in the CJR and control groups:

- 77 mandatory CJR participant hospitals and 33 control hospitals under the disproportionate share percentage definition
- 64 mandatory CJR participant hospitals and 16 control hospitals under the dually eligible percentage definition
- 69 mandatory CJR participant hospitals and 87 control hospitals under the ADI definition

Overlap between the disproportionate share percentage and dually eligible percentage definitions was high: 60% of hospitals that qualified as SNHs with the disproportionate share percentage

[62] Disproportionate share percentage is based on Fiscal Year (FY) 2022 public use data in the Inpatient Prospective Payment System Final Rule and Correction Notice files. Dual eligibility percentage is based on the FY22 Supplemental Case-Mix data file. ADI is based on claims data.

definition also qualified as an SNH using the dually eligible percentage definition. The ADI-based SNH definition had low overlap with both the disproportionate share percentage and dually eligible percentage definitions (**Exhibit V-4 and Exhibit V-5**).

Exhibit V-4: More Hospitals Satisfied the ADI Definition and Disproportionate Share Percentage Definition Than the Dually Eligible Percentage Definition

SNH Definition	Number of hospitals in sample	% of all hospitals in sample
Disproportionate share percentage	110	17.4%
Dually eligible percentage	80	12.7%
ADI	156	24.7%

Source: CJR evaluation team analysis of Medicare claims and enrollment data for episodes ending between October 1, 2021, and December 31, 2022 (PY 6 intervention), and the Fiscal Year 2022 Inpatient Prospective Payment System Final Rule and Correction Notice and Supplemental Case-Mix files.

Notes: Hospitals satisfying the Disproportionate Share Percentage definition were in the top 15% of hospitals nationwide for Disproportionate Share Percentage. Hospitals satisfying the dually eligible percentage definition had at least 40% of all hospital patients in FY22 having dual eligibility for Medicare and Medicaid. Hospitals satisfying the ADI definition had at least 30% of PY 6 LEJR patients residing in a 9-digit ZIP Code with ADI greater than 70. ADI = Area Deprivation Index; LEJR= lower extremity joint replacement; SNH = safety-net hospital.

Exhibit V-5: Hospitals That Satisfied the Disproportionate Share Percentage SNH Definition Often Also Satisfied the Dually Eligible Definition, but Hospitals That Satisfied the ADI SNH Definition Rarely Satisfied Another Definition

Definition combinations	Number of hospitals meeting multiple definitions
Disproportionate share percentage and dually eligible percentage	66
Disproportionate share percentage and ADI	20
Dually eligible percentage and ADI	4
Disproportionate share percentage, dually eligible percentage, and ADI	2

Source: CJR evaluation team analysis of Medicare claims and enrollment data for episodes ending between October 1, 2021, and December 31, 2022 (PY 6 intervention), and Fiscal Year 2022 Inpatient Prospective Payment System Final Rule and Correction Notice and Supplemental Case-Mix files.

Notes: Hospitals satisfying the Disproportionate Share Percentage definition were in the top 15% of hospitals nationwide for Disproportionate Share Percentage. Hospitals satisfying the dually eligible percentage definition had at least 40% of all hospital patients in FY22 having dual eligibility for Medicare and Medicaid. Hospitals satisfying the ADI definition had at least 30% of PY 6 LEJR patients residing in a 9-digit ZIP Code with ADI greater than 70. ADI = Area Deprivation Index; LEJR= lower extremity joint replacement; SNH = safety-net hospital.

b. Sample Selection and Analysis

Prior quantitative findings suggested disparities between underserved populations and reference populations in volume of LEJRs, quality, and utilization.[63] Prior qualitative findings also identified nonmedical factors such as social context, physical environments, and economic stability as

[63] Additional discussion of the disparities between underserved and reference populations can be found in Centers for Medicare & Medicaid Services. (2023). *Comprehensive Care for Joint Replacement Model - Fifth Annual Report*. https://www.cms.gov/priorities/innovation/data-and-reports/2023/cjr-py5-annual-report

influences in LEJR patient care pathways.[64] To expand on these previous findings, we conducted telephone interviews with CJR participant hospitals identified as SNHs. Interviews occurred between November and December 2023. We interviewed representatives from CJR participant hospitals including hospital staff and leadership, orthopedic surgeons, and system-level representatives. The interviews aimed to allow us to better understand how SNHs cared for LEJR patients with unmet nonmedical needs and how these hospitals interacted with the CJR Model. We categorized a hospital as a SNH if they were in the 85th percentile or higher in disproportionate share hospital patient percentage among hospitals nationally. We also accounted for dual-eligibility percentage, ADI, and geographic variability in selecting the sample.

We conducted outreach to 11 hospitals in our sample by sending an initial email invitation, asking if they would participate in this round of interviews. We attached a frequently asked questions document and informed consent information to the email. Once the hospital responded, we followed up with a confirmation email. Six hospitals of the eleven we conducted outreach to agreed to participate in our interviews. We did not continue outreach to the remaining hospitals in our sample.

Thematic analysis of telephone interview data was completed by notetakers and interviewers who participated in interviews. We developed analytic codebooks including primary and sub-codes based on the telephone interview protocols. Coders used ATLAS.ti to apply codes and sub-codes to comprehensive interview notes and ran queries to identify themes across interviews. All coders received systematic training, which included parallel coding and discussion of results with trainers until consistency was established. Throughout the analysis the codebooks were refined. Codes were dropped, consolidated, added, or revised to better capture patterns as they emerged.[65]

3. Results

All six hospitals interviewed had a low volume of CJR episodes (range 21–57 episodes) and made repayment to CMS under CJR (total net reconciliation payment amounts ranged from –$228K to –$921K).[66] Interviewees described their patient population as complex with high rates of underinsured or uninsured patients, comorbidities, dual eligibility among patients, and Medicare Advantage penetration in their market. Overall, interviewees from the six SNHs were not optimistic about financial prospects in the model due to low CJR episode volume and target prices that were "too low" or did not reflect the cost of care for their complex patient population.

Identifying and addressing SDOH was a key focus for the hospitals interviewed. In caring for LEJR patients with unmet non-medical needs, all six hospitals conducted screening of the social needs for all patients, utilizing standardized, electronically integrated SDOH screenings. Half of the hospitals continued this screening in their discharge planning.

---

[64] Additional discussion regarding qualitative findings for LEJR patient care pathways can be found in the Third Annual Report. Centers for Medicare & Medicaid Services. (2020). *Comprehensive Care for Joint Replacement Model - Third Annual Report*. https://www.cms.gov/priorities/innovation/data-and-reports/2020/cjr-thirdannrpt

[65] Full details on all the methods used for the analyses in this section can be found in **Appendix B: Data and Methods**.

[66] ***Source:*** CJR evaluation team analysis of Medicare claims and enrollment data for episodes ending in or between Quarter 4 (Q4) 2021–Q3 2022 (PY 6); CJR evaluation team analysis of Medicare claims for episodes ending Q4 2016–Q3 2021 (PY 1–5).

Lack of in-home family support, transportation, access to quality PAC services, access to regular primary care, and food security were common unmet non-medical needs of LEJR patients at SNHs.

Patients often required additional ongoing support such as help with treatment compliance and additional follow-up care not required with a less complex patient population. Language and cultural differences were also reported as barriers, specifically for access and acceptability of culturally appropriate PAC services. Specifically, hospitals serving patients who are Hispanic or Vietnamese reported hesitation in utilizing PAC services after receiving an LEJR when the provider did not speak their language and was geographically outside of their community.

In response to these barriers, the hospitals described how they implemented several improvements, such as additional investments in physical therapy equipment to improve treatment adherence post-discharge, increased partnerships with PAC providers to address language and cultural concerns, and further developments in community partnerships to provide resources such as primary care and educational support. However, hospitals stated that these improvements were implemented as a response to the needs of the patient population, independent of the CJR Model. Due to the overall low volume of Medicare FFS LEJR cases, the SNHs interviewed did not report introducing any changes due to the CJR Model.

Additionally, they reported that care transformation strategies that can improve financial success are often not feasible for SNHs. Although many hospitals can shift procedures to the outpatient setting to deliver care at a lower target price, as well as reduce the use of skilled nursing facilities (SNFs) and inpatient rehabilitation facilities (IRFs), SNHs said they could not make these changes because their patients' needs often require inpatient procedures and more post-discharge care.

To better provide care for LEJR patients with unmet nonmedical needs, these SNHs suggested additional funding for community services to address social barriers such as housing, transportation, mental and behavioral health services, and increased coverage for additional in-home care. They also requested financial improvements such as target pricing that reflects the needs of their complex patient populations and higher reimbursement for LEJR procedures to improve their financial performance under the CJR Model.

C. Did the CJR Model Impact the LEJR Volume for Underserved Populations?

During the baseline and intervention periods, underserved populations had higher average episode payments and used more institutional PAC relative to their corresponding reference populations.[67] Thus, hospitals could potentially reduce their average episode payments and increase reconciliation payments by providing fewer LEJRs to patients from underserved populations. In PY 6, changes in target pricing to better reflect patient characteristics likely changed the incentives for hospitals. Providing fewer LEJRs to patients from underserved populations would still reduce average episode payments but would only increase reconciliation payments to the extent that the new risk-adjustment on target prices does not capture the higher expected costs associated with underserved populations. In this section, we present analysis results on the rate of LEJRs conducted for the underserved populations we studied.

[67] Risk-adjusted baseline and PY 6 mean values for average episode payments and institutional PAC use can be found in **Appendix F: Claims-Based Health Equity Analyses**.

1. Summary of Findings

The callout box presents a summary of our findings. Results with an asterisk (*) are key findings.

- (*) The CJR Model did not have a substantial impact on LEJR volume for patients who are Black or African American or for patients who are Black or African American and dually eligible for Medicaid.
- (*) Patients who are Hispanic had 71 more LEJRs per 100,000 beneficiaries per year due to the CJR Model, and this finding was statistically significant.
- (*) Patients with dual eligibility had 41 more LEJRs per 100,000 beneficiaries per year due to the CJR Model, and this finding was nearly statistically significant.
- Underserved populations had substantially lower LEJR volume during the baseline and intervention period compared to their non-underserved reference populations. The CJR Model did not substantially affect the observed gap in LEJR volume between the underserved populations compared to their non-underserved reference populations.
- Estimated changes in LEJR volume differed for patients who are Black or African American or Hispanic relative to patients who are Non-Hispanic White but not enough to achieve statistical significance.

2. Methods

We estimated the impact of the CJR Model on the probability of receiving at least one elective LEJR for underserved populations and their non-underserved reference populations. We used a difference-in-difference-in-differences (DDD) approach, which estimated the impact of the CJR Model on two populations and simultaneously estimated the difference between the estimated impacts for these populations. This approach produced three estimates: a CJR impact estimate for the underserved population, a CJR impact estimate for the reference population, and an estimate of the "differential impact" of the CJR Model.

For this analysis, the data included all Medicare FFS beneficiaries living in mandatory CJR MSAs or their corresponding control MSAs between 2012 and 2022, excluding beneficiaries eligible for Medicare based on end-stage renal disease status. We excluded Calendar Years 2016–2021 from the analysis to focus on the impact of the model in 2022. PY 6 constitutes a 5-quarter period, but this analysis captured only the last 4 of those 5 quarters.

For the LEJR volume analysis, we included *all* elective LEJRs, regardless of the location of the procedure. Thus, the included LEJRs encompassed LEJRs that took place in the inpatient and outpatient setting, regardless of hospital participation, as well as LEJRs performed in ambulatory surgical centers.[68]

We considered only elective LEJRs because we are testing for discretion and structural features that may create barriers to LEJRs among underserved populations. Hip fractures present hospitals with

[68] Recall that some hospitals located in mandatory MSAs are not participants in PY 6 due to low-volume or rural designation.

limited opportunities for presurgery patient optimization or discretion in scheduling of the surgery. A hip replacement is the main clinical response to an emergent hip fracture, so there is little reason to believe *ex ante* that hospitals could have an impact on the rate or volume of fracture LEJRs, even if incentives to do so are present.

The analysis generated estimates in terms of the probability of a Medicare FFS beneficiary receiving at least one elective LEJR in a given year.[69] For ease of exposition, we refer to and interpret these results as LEJR rates per 100,000 FFS beneficiaries per year.[70]

3. Results

All four of the underserved populations studied had substantially lower rates of LEJRs in the baseline period than their reference populations, ranging from 42.5% to 58.9% of the baseline LEJR rate for the reference population (**Exhibit V-6**).[71] The CJR Model did not substantially affect the observed gap in LEJR volume between the underserved populations compared to their -non-underserved reference populations. These differences were risk-adjusted to account for variation in beneficiary demographics, clinical characteristics, and geographic indicators.

Exhibit V-6: In Both CJR and Control MSAs, Underserved Populations Had Substantially Lower Risk-Adjusted LEJR Rates Than the Reference Populations During the Baseline

Underserved population	Reference population	Baseline difference in CJR MSAs (LEJRs per 100k beneficiary-years)	Difference as % of reference population rate	Baseline difference in control MSAs (LEJRs per 100k beneficiary-years)	Difference as % of reference population rate
Beneficiaries who are Black or African American	Beneficiaries who are Non-Hispanic White [a]	-508	42.5%	-555	44.8%
Beneficiaries who were dually eligible	Beneficiaries who were not dually eligible	-686	56.4%	-662	52.9%
Beneficiaries who are Black or African American and dually eligible	Beneficiaries who are Non-Hispanic White and not dually eligible	-737	58.9%	-758	58.4%
Beneficiaries who are Hispanic	Beneficiaries who are Non-Hispanic White [a]	-549	46.0%	-549	44.5%

Source: CJR evaluation team analysis of Medicare claims and enrollment data for Medicare FFS beneficiaries enrolled in Medicare between 2012 and 2015 (baseline) or during 2022 (intervention).

[69] Approximately 4% of beneficiaries in CJR and control MSAs in 2022 had more than one LEJR.
[70] Full details on all the methods used for the analyses in this section can be found in **Appendix B: Data and Methods**.
[71] Baseline and PY 6 risk-adjusted LEJR rates can be found in **Appendix F: Claims-Based Health Equity Analyses, Exhibit F-1A**.

Notes: Negative differences indicate that the underserved population had a lower risk-adjusted baseline LEJR rate than the reference population. CJR = Comprehensive Care for Joint Replacement; k = thousand; LEJR = lower extremity joint replacement; MSA = metropolitan statistical area.

[a] The estimate for this reference population differs depending on the underserved population because each underserved-reference population pair is handled in a separate regression.

The CJR Model had noticeable impacts on two of the four underserved populations (**Exhibit V-7**). Beneficiaries who were dually eligible had an estimated increase in their LEJR rate of 41 LEJRs per 100,000 beneficiary-years, a 7.6% increase relative to their risk-adjusted baseline rate (p=0.11).[72] The estimate was nearly statistically significant at the 10% level. Relative to beneficiaries who were not dually eligible, the differential impact for beneficiaries who were dually eligible was a decrease of two LEJRs per 100,000 beneficiary-years (p=0.94), which was not statistically significant.

Exhibit V-7: The CJR Model Had a Statistically Significant Impact on LEJR Volume for Beneficiaries Who Are Hispanic

Population	DiD	p-value (DiD)	90% LCI (DiD)	90% UCI (DiD)	DDD (differential impact)	p-value (DDD)	90% LCI (DDD)	90% UCI (DDD)
Beneficiaries who are Black or African American	-11	0.77	-72	50	-56	0.13	-116	5
Beneficiaries who are Non-Hispanic White [a]	45	0.15	-6	95				
Beneficiaries who were dually eligible	41	0.11	-1	82	-2	0.94	-41	37
Beneficiaries who were not dually eligible	42	0.15	-6	90				
Beneficiaries who are Black or African American and dually eligible	-1	0.98	-76	74	-47	0.36	-132	38
Beneficiaries who are Non-Hispanic White and not dually eligible	46	0.11	-2	94				
Beneficiaries who are Hispanic	71	<0.01	28	115	28	0.40	-26	82
Beneficiaries who are Non-Hispanic White [a]	43	0.16	-7	94				

Source: CJR evaluation team analysis of Medicare claims and enrollment data for Medicare FFS beneficiaries enrolled in Medicare between 2012 and 2015 (baseline) or during 2022 (intervention).

Notes: The estimates in this exhibit are the result of a DDD model. DDD estimates that are significant at the 1%, 5%, or 10% significance levels are indicated by red, orange, or yellow shading, respectively. A full table of results, including sample counts and risk-adjusted means, is located in **Appendix F: Claims-Based Health Equity Analyses, Exhibits F-1A and F-1B**, and outcome definitions are in **Appendix B: Data and Methods, Exhibit B-6**. DiD = difference-in-differences; DDD = difference-in-difference-in-differences; LCI = lower confidence interval; UCI = upper confidence interval.

[a] The estimate for this reference population differs depending on the underserved population because each underserved-reference population pair is handled in a separate regression.

[72] Where we report percentages, they are calculated using the precise estimated values. There may be minor discrepancies between reported percentages and percentages calculated using the rounded numbers.

We estimated an increase in the LEJR rate for beneficiaries who are Hispanic of 71 LEJRs per 100,000 beneficiary-years, an 11.1% increase relative to their risk-adjusted baseline rate, which was statistically significant at the 1% level. Relative to beneficiaries who are Non-Hispanic White, beneficiaries who are Hispanic experienced an additional increase of 28 LEJRs per 100,000 beneficiary-years (p=0.40), which reduced the risk-adjusted difference in LEJR rates by 5.1%, although this differential impact was not statistically significant.

Beneficiaries who are Black or African American had an estimated decrease in the LEJR rate of 11 LEJRs per 100,000 beneficiary-years, a decrease of 1.6% relative to their risk-adjusted baseline rate (p=0.77). Relative to beneficiaries who are Non-Hispanic White, this was a decrease of 56 LEJRs per 100,000 beneficiary years, widening the difference in risk-adjusted LEJR rates by 10.9% (p=0.13). Although neither estimate was statistically significant, the differential impact had a p-value that is close to common statistical significance thresholds and indicated a possible widening of the existing baseline difference in LEJR rates between these populations.

Beneficiaries who are Black or African American and dually eligible had an estimated decrease in LEJR rates of 1 LEJR per 100,000 beneficiary-years (p=0.98), while beneficiaries who are Non-Hispanic White had an estimated increase of 46 LEJRs per 100,000 beneficiary-years, an increase of 3.7% relative to the baseline risk-adjusted rate (p=0.11). Thus, the estimated decrease for beneficiaries who are Black or African American and dually eligible relative to beneficiaries who are Non-Hispanic White was 47 LEJRs per 100,000 beneficiary-years, widening the difference in risk-adjusted LEJR rates by 6.4%, but this estimate was not close to statistical significance (p=0.36).

D. Did the CJR Model Impact the LEJR Cost, Utilization, and Quality Outcomes for Underserved Populations in PY 6?

This section provides subgroup impact analyses for underserved populations using data from PY 6. We studied three groups of claims outcomes: average episode payments, first PAC discharge destination, and PAC utilization measures.

1. Summary of Findings

- (*) The CJR Model likely reduced average episode payments for all the studied populations. Estimated reductions were as large or greater for underserved populations relative to their respective reference populations.
- (*) The estimated payment reductions for all underserved populations were primarily driven by changes in PAC discharge destination, the same driver of estimated payment reductions for all populations.
- (*) Evidence suggests that CJR increased ED use for underserved populations.
- (*) Underserved populations were relatively more likely to be discharged home without home health than reference populations.
- (*) We did not observe notable changes in patient characteristics across the underserved populations studied between the baseline period and PY 6.

2. Methods

To analyze the impact of the CJR Model on payments, utilization, and quality for underserved populations, we again used a DDD approach, as we did in the volume analysis above. For these analyses, we used the research file comprising LEJR episodes at CJR and control hospitals during the baseline period and PY 6. We studied the following claims-based outcomes:

- Average episode payments
- PAC utilization measures:
  - First PAC discharge destination
  - SNF days and IRF days in the 90 days post-discharge
  - Home health visits in the 90 days post-discharge
- Quality Measures:
  - All-cause mortality during the hospital stay or in the 90 days post-discharge
  - ED use in the 90 days post-discharge
  - Unplanned readmission rate during the 90 days post-discharge

Similar to the prior analysis focused on LEJR rates, the DDD approach for cost, utilization, and quality outcomes generated three estimates for each outcome: an estimated model impact for the underserved population, an estimated model impact for the reference population, and an estimate of the difference between the impacts (the differential impact). The differential impact is interpreted as the impact of the CJR Model on the baseline differences in risk-adjusted outcome levels between the two populations.

We used the same claims-based dataset that we used for our main impact analyses, including episode eligibility restrictions. We dropped PYs 1–5 from the analysis to focus on the impact of the CJR Model in PY 6.[73]

Additionally, we considered only elective LEJRs. We observed that model impacts for fracture and elective LEJRs differed substantially, sometimes even pointing in different directions. For this reason, we decided to isolate the impact on underserved populations for elective LEJRs. We could not produce fracture LEJR estimates due to sample size. For additional information on CJR Model impacts, see **Chapter III: Impact of the Model**.

Exhibit V-8 provides the sample sizes for the CJR and control groups' elective LEJR episodes in the seven populations included in this analysis (four underserved populations and three reference populations; patients who are Non-Hispanic White were used as a reference population twice). Note that results for patients who are both Black or African American and dually eligible pertain to a small population.

[73] Full details on all the methods used for the analyses in this section can be found in **Appendix B: Data and Methods**.

Exhibit V-8: Sample Sizes for Average Episode Payments Differential Impact Analysis

Population	CJR PY 6 sample	Control PY 6 sample	CJR baseline sample	Control baseline sample	Total sample size
Patients who are Black or African American	1,943	2,662	6,181	8,907	19,693
Patients who were dually eligible	3,103	2,577	12,038	12,017	29,735
Patients who are Black or African American and dually eligible	396	483	2,381	3,061	6,321
Patients who are Hispanic	2,312	1,328	6,195	3,676	13,511
Patients who are Non-Hispanic White	40,046	44,909	80,932	107,735	273,622
Patients who were not dually eligible	43,908	48,502	84,228	110,517	287,155
Patients who are Non-Hispanic White and not dually eligible	38,597	43,398	75,695	100,856	258,546

Source: CJR evaluation team analysis of Medicare claims and enrollment data for episodes initiated between January 1, 2012, and December 31, 2014, that ended between April 1, 2012, and March 31, 2015 (baseline) and episodes that ended between October 1, 2021, and December 31, 2022 (PY 6 intervention).

Notes: Sample sizes for average episode payments DDD analysis. Sample sizes for other outcomes are smaller. A full table of results, including sample counts and risk-adjusted means, is in **Appendix F: Claims-Based Health Equity Analyses**. CJR = Comprehensive Care for Joint Replacement; DDD = difference-in-difference-in-differences; PY = performance year.

3. Results

a. Differential Impact for Average Episode Payments

The CJR Model likely reduced average episode payments for all seven of the studied populations (**Exhibit V-9**). Patients who are Black or African American or Hispanic had an estimated reduction in average episode payments of $2,132, approximately twice as large as that for patients who are Non-Hispanic White. Patients who were dually eligible had an estimated reduction of $2,008, almost twice as large as that for patients who were not dually eligible. Patients who are Hispanic had an estimated reduction of $2,191, almost twice as large as that for patients who are Non-Hispanic White. All four underserved populations had higher average episode payments in the baseline and continued to have higher average episode payments in PY 6, but the difference between the underserved populations and their reference populations likely shrank somewhat.

Exhibit V-9: The CJR Model Likely Reduced Average Episode Payments for All Patients, but Reduced Average Episode Payments by More for Patients From Underserved Populations

Population	DiD (impact)	p-value (DiD)	90% LCI (DiD)	90% UCI (DiD)	DDD (differential impact)	p-value (DDD)	90% LCI (DDD)	90% UCI (DDD)
Patients who are Black or African American	-$2,132[a]	0.07	-$4,056	-$207	-$980[a]	0.15	-$2,105	$146
Patients who are Non-Hispanic White[b]	-$1,152	0.10	-$2,296	-$8				

Population	DiD (impact)	p-value (DiD)	90% LCI (DiD)	90% UCI (DiD)	DDD (differential impact)	p-value (DDD)	90% LCI (DDD)	90% UCI (DDD)
Patients who were dually eligible	-$2,008[a]	0.02	-$3,431	-$585	-$847	0.08	-$1,635	-$58
Patients who were not dually eligible	-$1,161	0.10	-$2,327	$5				
Patients who are Black or African American and dually eligible	-$1,980[a]	0.24	-$4,777	$817	-$883	0.47	-$2,921	$1,154
Patients who are Non-Hispanic White and not dually eligible	-$1,097	0.11	-$2,236	$42				
Patients who are Hispanic	-$2,191	<0.01	-$3,469	-$913	-$1,051	0.01	-$1,739	-$363
Patients who are Non-Hispanic White [b]	-$1,140	0.10	-$2,289	$9				

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 ending between October 1, 2021, and December 31, 2022 (PY 6 intervention).

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

A full table of results, including sample counts and risk-adjusted means, is located in **Appendix F: Claims-Based Health Equity Analyses, Exhibits F-2A and F-2B**, and outcome definitions are in **Appendix B: Data and Methods, Exhibit B-6**. DiD = difference-in-differences; DDD = difference-in-difference-in-differences; LCI = lower confidence interval; UCI = upper confidence interval.

[a] The sample for this estimate failed tests for parallel trends in the baseline period.

[b] The estimate for this reference population differs depending on the underserved population because each underserved-reference population pair is handled in a separate regression.

b. Differential Impact for Quality and PAC Utilization

For the remaining results, which cover quality and PAC utilization measures, we report results separated by the populations studied.

Patients Who Are Black or African American

The CJR Model reduced the likelihood of discharge to IRFs for patients who are Black or African American and patients who are Non-Hispanic White by 6.5 percentage points and 3.3 percentage points respectively (**Exhibit V-10**). For discharge with home health, patients who are Black or African American and patients who are Non-Hispanic White had estimated increases in PAC utilization of 7 and 11.8 percentage points respectively. For discharge home without home health, the estimated impacts were an increase for patients who are Black or African American and a decrease for patients who are Non-Hispanic White. While neither impact alone was statistically significant, the difference between them was.

Exhibit V-10: The CJR Model Led to a Larger Decrease in First Discharge to IRF for Patients Who Are Black or African American Than for Patients Who Are Non-Hispanic White and a Significantly Different Impact on First Discharge Without Home Health

First PAC discharge	Population	DiD (impact)	p-value (DiD)	90% LCI (DiD)	90% UCI (DiD)	DDD (differential impact)	p-value (DDD)	90% LCI (DDD)	90% UCI (DDD)
IRF	Patients who are Black or African American	-6.5 pp	0.03	-11.3 pp	-1.6 pp	-3.1 pp	0.09	-6.2 pp	-0.1 pp
	Patients who are Non-Hispanic White	-3.3 pp	0.10	-6.6 pp	0.0 pp				
SNF	Patients who are Black or African American	-2.0 pp	0.65	-9.3 pp	5.3 pp	1.4 pp	0.66	-3.9 pp	6.7 pp
	Patients who are Non-Hispanic White	-3.4 pp	0.14	-7.1 pp	0.4 pp				
Home health	Patients who are Black or African American	7.0 pp	0.42	-7.2 pp	21.1 pp	-4.8 pp	0.24	-11.6 pp	2.0 pp
	Patients who are Non-Hispanic White	11.8 pp[a]	0.07	1.0 pp	22.5 pp				
No home health	Patients who are Black or African American	1.5 pp	0.78	-7.1 pp	10.1 pp	6.5 pp	<0.01	2.9 pp	10.1 pp
	Patients who are Non-Hispanic White	-5.1 pp	0.28	-12.8 pp	2.7 pp				

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 ending between October 1, 2021, and December 31, 2022 (PY 6 intervention).

Notes: The estimates in this exhibit are the result of a DDD model. DDD estimates that are significant at the 1%, 5%, or 10% significance levels are indicated by red, orange, or yellow shading, respectively. A full table of results, including sample counts and risk-adjusted means, is in **Appendix F: Claims-Based Health Equity Analyses, Exhibits F-4A and F-4B**, and outcome definitions are in **Appendix B: Data and Methods, Exhibit B-6**. CJR = Comprehensive Care for Joint Replacement; DiD = difference-in-differences; DDD = difference-in-difference-in-differences; IRF = Inpatient rehabilitation facility; LCI = lower confidence interval; pp = percentage point; SNF = skilled nursing facility; UCI = upper confidence interval.

[a] The sample for this estimate failed tests for parallel trends in the baseline period.

For both the Black or African American and Non-Hispanic White populations, the CJR Model resulted in a statistically significant decrease in first discharge to an IRF. This decrease was greater for patients who are Black or African American than for patients who are Non-Hispanic White, with a statistically significant differential impact of –3.1 percentage points (p=0.09). Both of these populations had statistically insignificant estimated decreases in first discharge to SNF. Patients who are Black or African American had a statistically insignificant estimated increase in first discharge to home with home health of 7.0 percentage points (p=0.42), while patients who are Non-Hispanic White had a larger estimated increase of 11.8 percentage points (p=0.07). The differential impact of –4.8 percentage points was not statistically significant (p=0.24). Although neither population had a

statistically significant estimated impact for first discharge to home without home health, the differential impact of 6.5 percentage points was statistically significant (p<0.01).

Taken together, these results suggest that patients who are Black or African American were relatively more likely than before the model to be discharged home without home health and relatively less likely to be discharged to an IRF compared with patients who are Non-Hispanic White.

The estimated impact on SNF length of stay for patients who are Black or African American was a decrease in average SNF length of stay of 3.4 days (p=0.06), with a nearly significant differential impact of –2.8 days (p=0.11) (**Exhibit V-11**). However, populations for both estimates failed tests for parallel trends in the baseline period, so we cannot confidently assign causality for these impacts to the CJR Model.

Neither population experienced a statistically significant change in average home health visits or IRF length of stay, and neither differential impact was statistically significant.

Exhibit V-11: Patients Who Are Black or African American Had a Statistically Significant Decrease in SNF Length of Stay, but the Difference Relative to Patients Who Are Non-Hispanic White Was Not Statistically Significant

PAC utilization	Population	DiD (impact)	p-value (DiD)	90% LCI (DiD)	90% UCI (DiD)	DDD (differential impact)	p-value (DDD)	90% LCI (DDD)	90% UCI (DDD)
HH visits	Patients who are Black or African American	-1.0	0.10	-2.0	0.0	-0.3	0.54	-1.1	0.5
	Patients who are Non-Hispanic White	-0.7[a]	0.24	-1.7	0.3				
SNF LOS	Patients who are Black or African American	-3.4[a]	0.06	-6.3	-0.5	-2.8[a]	0.11	-5.7	0.0
	Patients who are Non-Hispanic White	-0.6	0.49	-2.0	0.8				
IRF LOS	Patients who are Black or African American	-0.4	0.65	-1.8	1.0	-0.8	0.35	-2.1	0.6
	Patients who are Non-Hispanic White	0.4	0.16	-0.1	0.8				

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 ending between October 1, 2021, and December 31, 2022 (PY 6 intervention).

Notes: The estimates in this exhibit are the result of a DDD model. DDD estimates that are significant at the 1%, 5%, or 10% significance levels are indicated by red, orange, or yellow shading, respectively. A full table of results, including sample counts and risk-adjusted means, is in **Appendix F: Claims-Based Health Equity Analyses, Exhibits F-5A and F-5B**, and outcome definitions are in **Appendix B: Data and Methods, Exhibit B-6**. CJR = Comprehensive Care for Joint Replacement; DiD = difference-in-differences; DDD = difference-in-difference-in-differences; HH = home health; IRF = Inpatient rehabilitation facility; LCI = lower confidence interval; pp = percentage point; SNF = skilled nursing facility; UCI = upper confidence interval.

[a] The sample for this estimate failed tests for parallel trends in the baseline period.

The estimated impact of the model on all-cause mortality for patients who are Black or African American was a decrease of 0.4 percentage points, which was statistically significant (p=0.08) (**Exhibit V-12**). Patients who are Non-Hispanic White also had an estimated decrease, but it was less than one-tenth of a percentage point and not statistically significant. The differential impact of –0.4 percentage points was nearly significant (p=0.11).

For ED use, patients who are Black or African American had a statistically significant estimated increase of 2.2 percentage points (p=0.08), while patients who are Non-Hispanic White had an estimated increase of less than one-tenth of a percentage point that was not statistically significant. The differential impact was 2.2 percentage points and was statistically significant (p=0.09).

Both populations had estimated decreases in the unplanned readmission rate. Patients who are Black or African American had a larger estimated decrease of 0.8 percentage points, but this reduction was not statistically significant (p=0.38); patients who are Non-Hispanic White had a statistically significant decrease of 0.5 percentage points (p=0.08). The differential impact was small and not statistically significant.

Exhibit V-12: Patients Who Are Black or African American Had a Statistically Significant Decrease in Mortality Rate and a Statistically Significant Increase in ED Use, Which Was Significantly Different From the Impact on ED Use for Patients Who Are Non-Hispanic White

Measure	Population	DiD (impact)	p-value (DiD)	90% LCI (DiD)	90% UCI (DiD)	DDD (differential impact)	p-value (DDD)	90% LCI (DDD)	90% UCI (DDD)
Mortality rate	Patients who are Black or African American	-0.4 pp	0.08	-0.8 pp	-0.0 pp	-0.4 pp	0.11	-0.8 pp	0.0 pp
	Patients who are Non-Hispanic White	0.0 pp	0.77	-0.1 pp	0.1 pp				
ED use	Patients who are Black or African American	2.2 pp	0.08	0.1 pp	4.3 pp	2.2 pp	0.09	0.1 pp	4.3 pp
	Patients who are Non-Hispanic White	0.0 pp	0.96	-0.7 pp	0.7 pp				
Readmission rate	Patients who are Black or African American	-0.8 pp	0.38	-2.3 pp	0.7 pp	-0.3 pp	0.76	-1.7 pp	1.2 pp
	Patients who are Non-Hispanic White	-0.5 pp	0.08	-1.0 pp	-0.0 pp				

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 ending between October 1, 2021, and December 31, 2022 (PY 6 intervention).

Notes: The estimates in this exhibit are the result of a DDD model. DDD estimates that are significant at the 1%, 5%, or 10% significance levels are indicated by red, orange, or yellow shading, respectively. A full table of results, including sample counts

and risk-adjusted means, is in **Appendix F: Claims-Based Health Equity Analyses, Exhibits F-3A and F-3B,** and outcome definitions are in **Appendix B: Data and Methods, Exhibit B-6**. CJR = Comprehensive Care for Joint Replacement; DiD = difference-in-differences; DDD = difference-in-difference-in-differences; ED = emergency department; LCI = lower confidence interval; pp = percentage point; UCI = upper confidence interval.

Patients Who Were Dually Eligible for Medicaid

Patients who were dually eligible had estimated decreases in IRF and SNF discharge rates that were slightly larger than the decreases for patients who were not dually eligible, but neither differential impact was statistically significant (**Exhibit V-13**). The model led to a statistically significant decrease in the IRF discharge rate of 3.5 percentage points (p=0.06) for patients who were not dually eligible. The decrease for patients who were dually eligible was larger but failed to achieve statistical significance, likely due to sample size.

Both patients who were and were not dually eligible had large and statistically significant increases for discharge to home with home health: Patients who were dually eligible had an estimated increase of 10.2 percentage points (p=0.05), and patients who were not dually eligible had an estimated increase of 11.2 percentage points (p=0.09). The differential impact was small and not statistically significant.

Although none of the impacts for discharge to home without home health were statistically significant, they pointed in the direction of a relative increase in discharge to home without home health for patients who were dually eligible compared with patients who were not dually eligible.

Exhibit V-13: Patients Who Were Dually Eligible Had a Statistically Significant Increase in First Discharge With Home Health, but No Significantly Different Differences Existed Relative to Patients Who Were Not Dually Eligible

First PAC discharge	Population	DiD (impact)	p-value (DiD)	90% LCI (DiD)	90% UCI (DiD)	DDD (Differential impact)	p-value (DDD)	90% LCI (DDD)	90% UCI (DDD)
IRF	Patients who were dually eligible	-4.1 pp	0.13	-8.6 pp	0.4 pp	-0.6 pp	0.69	-3.2 pp	1.9 pp
	Patients who were not dually eligible	-3.5 pp	0.06	-6.6 pp	-0.4 pp				
SNF	Patients who were dually eligible	-4.8 pp	0.13	-10.0 pp	0.4 pp	-1.4 pp	0.46	-4.7 pp	1.8 pp
	Patients who were not dually eligible	-3.4 pp	0.15	-7.2 pp	0.4 pp				
Home health	Patients who were dually eligible	10.2 pp	0.05	1.7 pp	18.6 pp	-1.1 pp	0.76	-6.9 pp	4.7 pp
	Patients who were not dually eligible	11.2 pp[a]	0.09	0.5 pp	22.0 pp				

First PAC discharge	Population	DiD (impact)	p-value (DiD)	90% LCI (DiD)	90% UCI (DiD)	DDD (Differential impact)	p-value (DDD)	90% LCI (DDD)	90% UCI (DDD)
No home health	Patients who were dually eligible	-1.2 pp	0.64	-5.5 pp	3.1 pp	3.2 pp	0.27	-1.6 pp	7.9 pp
	Patients who were not dually eligible	-4.4 pp	0.35	-12.2 pp	3.4 pp				

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 ending between October 1, 2021, and December 31, 2022 (PY 6 intervention).

Notes: The estimates in this exhibit are the result of a DDD model. DDD estimates that are significant at the 1%, 5%, or 10% significance levels are indicated by red, orange, or yellow shading, respectively. A full table of results, including sample counts and risk-adjusted means, is located in **Appendix F: Claims-Based Health Equity Analyses, Exhibits F-7A and F-7B**, and outcome definitions are in **Appendix B: Data and Methods, Exhibit B-6**. CJR = Comprehensive Care for Joint Replacement; DiD = difference-in-differences; DDD = difference-in-difference-in-differences; IRF = inpatient rehabilitation facility; LCI = lower confidence interval; pp = percentage point; PY = performance year; SNF = skilled nursing facility; UCI = upper confidence interval.

[a] The sample for this estimate failed tests for parallel trends in the baseline period.

Patients who were and were not dually eligible had similar estimated impacts for the average number of home health visits, although the populations for both impact estimates failed parallel pre-trend tests (**Exhibit V-14**). The differential impact was small and not significant.

Patients who were dually eligible had a much larger estimated impact for average SNF length of stay: a reduction of 2.8 days, which was statistically significant (p=0.02). The differential impact was a relative reduction of 2.2 days, which was also significant (p=0.08). Patients who were not dually eligible had a small and statistically insignificant estimated reduction of 0.6 days.

Patients who were not dually eligible had a statistically significant estimated impact for IRF length of stay—an increase of 0.5 days (p=0.07)—while patients who were dually eligible had an estimated decrease of 0.2 days (p=0.81). The differential impact was not significant.

Exhibit V-14: Patients Who Were Dually Eligible Had a Statistically Significant Decrease in SNF Length of Stay and the Difference Relative to Patients Who Were Not Dually Eligible Was Statistically Significant

PAC utilization	Population	DiD (impact)	p-value (DiD)	90% LCI (DiD)	90% UCI (DiD)	DDD (Differential impact)	p-value (DDD)	90% LCI (DDD)	90% UCI (DDD)
HH visits	Patients who were dually eligible	-0.5[a]	0.58	-1.9	0.9	0.3	0.78	-1.3	1.8
	Patients who were not dually eligible	-0.7[a]	0.24	-1.8	0.3				
SNF LOS	Patients who were dually eligible	-2.8[a]	0.02	-4.9	-0.8	-2.2	0.08	-4.3	-0.2
	Patients who were not dually eligible	-0.6	0.49	-2.0	0.8				
IRF LOS	Patients who were dually eligible	-0.2	0.81	-1.5	1.1	-0.7	0.36	-1.9	0.6
	Patients who were not dually eligible	0.5	0.07	0.0	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 ending between October 1, 2021, and December 31, 2022 (PY 6 intervention).

Notes: The estimates in this exhibit are the result of a DDD model. DDD estimates that are significant at the 1%, 5%, or 10% significance levels are indicated by red, orange, or yellow shading, respectively. A full table of results, including sample counts and risk-adjusted means, is in **Appendix F: Claims-Based Health Equity Analyses, Exhibits F-8A and F-8B**, and outcome definitions are in **Appendix B: Data and Methods, Exhibit B-6**. CJR = Comprehensive Care for Joint Replacement; DiD = difference-in-differences; DDD = difference-in-difference-in-differences; HH = home health; IRF = inpatient rehabilitation facility; LCI = lower confidence interval; pp = percentage point; SNF = skilled nursing facility; UCI = upper confidence interval.

[a] The sample for this estimate failed tests for parallel trends in the baseline period.

Patients who were and were not dually eligible both had small decreases in all-cause mortality and unplanned readmission rates that were not statistically significant, and differential impacts between the populations were statistical zeros (**Exhibit V-15**). Patients who were dually eligible, however, had a moderately large increase in ED use of 2.1 percentage points, but it was not statistically significant (p=0.14). Similar to the results for patients who are Black or African American relative to patients who are Non-Hispanic White, the model led to a greater increase in ED use for patients who were dually eligible relative to patients who were not dually eligible; however, the differential impact between patients who were dually eligible and those who were not dually eligible was not statistically significant.

Exhibit V-15: The Model Did Not Have Any Significant Impacts on Quality Measures for Patients Who Were or Were Not Dually Eligible, and Differences Between the Populations Were Not Statistically Significant

Measure	Population	DiD (impact)	p-value (DiD)	90% LCI (DiD)	90% UCI (DiD)	DDD (differential impact)	p-value (DDD)	90% LCI (DDD)	90% UCI (DDD)
Mortality rate	Patients who were dually eligible	-0.1 pp	0.80	-0.5 pp	0.4 pp	0.0 pp	0.88	-0.5 pp	0.4 pp
	Patients who were not dually eligible	0.0 pp	0.68	-0.1 pp	0.1 pp				
ED use	Patients who were dually eligible	2.1 pp	0.14	-0.2 pp	4.4 pp	1.8 pp	0.16	-0.3 pp	4.0 pp
	Patients who were not dually eligible	0.2 pp	0.53	-0.4 pp	0.9 pp				
Readmission Rate	Patients who were dually eligible	-0.5 pp	0.39	-1.5 pp	0.5 pp	-0.1 pp	0.91	-1.1 pp	1.0 pp
	Patients who were not dually eligible	-0.5 pp	0.14	-1.0 pp	0.1 pp				

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 ending between October 1, 2021, and December 31, 2022 (PY 6 intervention).

Notes: The estimates in this exhibit are the result of a DDD model. DDD estimates that are significant at the 1%, 5%, or 10% significance levels are indicated by red, orange, or yellow shading, respectively. A full table of results, including sample counts and risk-adjusted means, is in **Appendix F: Claims-Based Health Equity Analyses, Exhibits F-6A and F-6B**, and outcome definitions are in **Appendix B: Data and Methods, Exhibit B-6**. CJR = Comprehensive Care for Joint Replacement; DiD = difference-in-differences; DDD = difference-in-difference-in-differences; ED = emergency department; LCI = lower confidence interval; pp = percentage point; UCI = upper confidence interval.

Patients Who Are Black or African American and Dually Eligible for Medicaid

Due to the small population of patients who are both Black or African American and dually eligible in PY 6, we did not observe statistically significant impacts on any measure for this population. Patients who are Black or African American and dually eligible and the reference population—patients who are Non-Hispanic White and not dually eligible—both had estimated decreases in IRF and SNF discharge rates (**Exhibit V-16**). Both populations also had estimated increases in discharge to home with home health, but the estimated increase was substantially larger for patients who are Non-Hispanic White and not dually eligible (11.8 percentage points) and was statistically significant (p=0.07). The differential impact of –8.0 percentage points, while large, was not statistically significant.

There was a statistically significant differential impact of 6.9 percentage points (p=0.02) for discharge to home without home health. While the small sample size limited our ability to make strong conclusions, this evidence again points toward a differential impact that suggested underserved populations were relatively more likely to be discharged to home without home health than their reference population.

Exhibit V-16: Patients Who Are Black or African American and Dually Eligible Had No Statistically Significant Impacts for First Discharge Destination, but the Differential Impact for Discharge Without Home Health Was Significant

First PAC discharge	Population	DiD (impact)	p-value (DiD)	90% LCI (DiD)	90% UCI (DiD)	DDD (differential impact)	p-value (DDD)	90% LCI (DDD)	90% UCI (DDD)
IRF	Patients who are Black or African American and dually eligible	-4.3 pp	0.15	-9.3 pp	0.6 pp	-1.0 pp[a]	0.63	-4.5 pp	2.5 pp
	Patients who are Non-Hispanic White and not dually eligible	-3.3 pp	0.08	-6.4 pp	-0.2 pp				
SNF	Patients who are Black or African American and dually eligible	-1.4 pp	0.79	-10.0 pp	7.3 pp	2.1 pp	0.62	-4.9 pp	9.1 pp
	Patients who are Non-Hispanic White and not dually eligible	-3.5 pp	0.14	-7.3 pp	0.4 pp				
Home health	Patients who are Black or African American and dually eligible	3.9 pp	0.65	-10.3 pp	18.1 pp	-8.0 pp	0.14	-16.9 pp	1.0 pp
	Patients who are Non-Hispanic White and not dually eligible	11.8 pp[a]	0.07	1.0 pp	22.7 pp				
No home health	Patients who are Black or African American and dually eligible	1.8 pp	0.73	-6.8 pp	10.5 pp	6.9 pp	0.02	1.9 pp	11.9 pp
	Patients who are Non-Hispanic White and not dually eligible	-5.1 pp	0.29	-13.0 pp	2.9 pp				

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 ending between October 1, 2021, and December 31, 2022 (PY 6 intervention).

Notes: The estimates in this exhibit are the result of a DDD model. DDD estimates that are significant at the 1%, 5%, or 10% significance levels are indicated by red, orange, or yellow shading, respectively. A full table of results, including sample counts and risk-adjusted means, is in **Appendix F: Claims-Based Health Equity Analyses, Exhibits F-10A and F-10B**, and outcome definitions are in **Appendix B: Data and Methods, Exhibit B-6**. CJR = Comprehensive Care for Joint Replacement; DiD = difference-in-differences; DDD = difference-in-difference-in-differences; IRF = inpatient rehabilitation facility; LCI = lower confidence interval; pp = percentage point; SNF = skilled nursing facility; UCI = upper confidence interval.

[a] The sample for this estimate failed tests for parallel trends in the baseline period.

We observed no statistically significant impacts, differential or otherwise, on PAC utilization measures for patients who are Black or African American and dually eligible (**Exhibit V-17**). The decrease in SNF length of stay was large for patients who are Black or African American and dually

eligible, and this decrease drove a large differential impact between this population and the reference population. However, the sample size issues weighed against drawing any substantial conclusions.

Exhibit V-17: Patients Who Are Black or African American and Dually Eligible Had No Statistically Significant Impacts for PAC Utilization Measures and No Significant Differences Compared With Patients Who Are Non-Hispanic White and Not Dually Eligible

PAC utilization	Population	DiD (impact)	p-value (DiD)	90% LCI (DiD)	90% UCI (DiD)	DDD (differential impact)	p-value (DDD)	90% LCI (DDD)	90% UCI (DDD)
HH visits	Patients who are Black or African American and dually eligible	-1.0	0.30	-2.7	0.6	-0.3	0.83	-2.4	1.8
	Patients who are Non-Hispanic White and not dually eligible	-0.8[a]	0.21	-1.8	0.3				
SNF LOS	Patients who are Black or African American and dually eligible	-4.6[a]	0.20	-10.6	1.4	-4.1	0.26	-10.2	1.9
	Patients who are Non-Hispanic White and not dually eligible	-0.5	0.52	-1.9	0.8				
IRF LOS	Patients who are Black or African American and dually eligible	-1.4	0.29	-3.7	0.8	-1.8	0.19	-4.1	0.4
	Patients who are Non-Hispanic White and not dually eligible	0.4	0.16	-0.1	0.8				

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 ending between October 1, 2021, and December 31, 2022 (PY 6 intervention).

Notes: The estimates in this exhibit are the result of a DDD model. DDD estimates that are significant at the 1%, 5%, or 10% significance levels are indicated by red, orange, or yellow shading, respectively. A full table of results, including sample counts and risk-adjusted means, is in **Appendix F: Claims-Based Health Equity Analyses, Exhibits F-11A and F-11B**, and outcome definitions are in **Appendix B: Data and Methods, Exhibit B-6**. CJR = Comprehensive Care for Joint Replacement; DiD = difference-in-differences; DDD = difference-in-difference-in-differences; HH = home health; IRF = inpatient rehabilitation facility; LCI = lower confidence interval; pp = percentage point; SNF = skilled nursing facility; UCI = upper confidence interval.

[a] The sample for this estimate failed tests for parallel trends in the baseline period.

Similarly, there was little evidence of quality impacts, differential or otherwise, on patients who are Black or African American and dually eligible. There was a notable estimated increase in ED use for patients who are Black or African American and dually eligible, which drove a substantial estimated differential impact between this population and the reference group, but neither estimate was statistically significant (**Exhibit V-18**). We again advise against drawing any substantial conclusions due to sample size concerns.

Exhibit V-18: Patients Who Are Black or African American and Dually Eligible Had No Significant Impacts for ED Use or Unplanned Readmission Rate, and No Significant Differences Relative to Patients Who Are Non-Hispanic White and Not Dually Eligible

Measure	Population	DiD (impact)	p-value (DiD)	90% LCI (DiD)	90% UCI (DiD)	DDD (differential impact)	p-value (DDD)	90% LCI (DDD)	90% UCI (DDD)
Mortality rate	Patients who are Black or African American and dually eligible	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A
	Patients who are Non-Hispanic White and not dually eligible	N/A	N/A	N/A	N/A				
ED use	Patients who are Black or African American and dually eligible	4.5 pp	0.14	-0.5 pp	9.4 pp	4.4 pp	0.16	-0.7 pp	9.6 pp
	Patients who are Non-Hispanic White and not dually eligible	0.0 pp	0.94	-0.7 pp	0.7 pp				
Readmission rate	Patients who are Black or African American and dually eligible	-0.4 pp	0.84	-3.3 pp	2.6 pp	0.1 pp	0.95	-2.7 pp	2.9 pp
	Patients who are Non-Hispanic White and not dually eligible	-0.5 pp	0.12	-1.0 pp	0.0 pp				

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 ending between October 1, 2021, and December 31, 2022 (PY 6 intervention).

Notes: The estimates in this exhibit are the result of a DDD model. DDD estimates that are significant at the 1%, 5%, or 10% significance levels are indicated by red, orange, or yellow shading, respectively. Estimates for All-Cause Mortality are not reported, as that regression is not identified due to zero CJR PY 6 episodes with a death during the episode for patients who are Black or African American and dually eligible. A full table of results, including sample counts and risk-adjusted means, is in **Appendix F: Claims-Based Health Equity Analyses, Exhibits F-9A and F-9B**, and outcome definitions are in **Appendix B: Data and Methods, Exhibit B-6**. CJR = Comprehensive Care for Joint Replacement; DiD = difference-in-differences; DDD = difference-in-difference-in-differences; ED = emergency department; LCI = lower confidence interval; pp = percentage point; UCI = upper confidence interval.

Patients Who Are Hispanic

Patients who are Hispanic and patients who are Non-Hispanic White had estimated decreases in first discharge to an IRF and SNF caused by the model (**Exhibit V-19**). Unlike the other populations above, patients who are Hispanic had a substantially larger estimated decrease in SNF discharge: a reduction of 8.3 percentage points (p=0.02), or an additional 4.8 percentage points (p=0.10) relative to patients who are Non-Hispanic White.

Both populations had positive and large estimated increases in discharge to home with home health, 13.6 percentage points (p=0.01) and 11.3 percentage points (p=0.08), respectively. The differential impact was not statistically significant.

Patients who are Hispanic had a statistically insignificant increase in discharges to home without home health, while patients who are Non-Hispanic White had a statistically insignificant decrease. The differential impact of 5.2 percentage points (p=0.11) was nearly statistically significant.

Exhibit V-19: Patients Who Are Hispanic Had a Significant Increase in First Discharge With Home Health and a Significant Decrease in First Discharge to SNF. Compared With Patients Who Are Non-Hispanic White, the Difference in Impact for First Discharge to SNF Was Significant

First PAC discharge	Population	DiD (impact)	p-value (DiD)	90% LCI (DiD)	90% UCI (DiD)	DDD (differential impact)	p-value (DDD)	90% LCI (DDD)	90% UCI (DDD)
IRF	Patients who are Hispanic	-5.8 pp	0.11	-11.7 pp	0.2 pp	-2.6 pp	0.47	-8.7 pp	3.4 pp
	Patients who are Non-Hispanic White	-3.1 pp	0.10	-6.2 pp	-0.0 pp				
SNF	Patients who are Hispanic	-8.3 pp	0.02	-14.2 pp	-2.3 pp	-4.8 pp	0.10	-9.6 pp	-0.0 pp
	Patients who are Non-Hispanic White	-3.5 pp	0.14	-7.4 pp	0.4 pp				
Home health	Patients who are Hispanic	13.6 pp	0.01	4.6 pp	22.6 pp	2.3 pp	0.53	-3.7 pp	8.2 pp
	Patients who are Non-Hispanic White	11.3 pp[a]	0.08	0.8 pp	21.9 pp				
No home health	Patients who are Hispanic	0.5 pp	0.88	-4.5 pp	5.4 pp	5.2 pp	0.11	-0.1 pp	10.4 pp
	Patients who are Non-Hispanic White	-4.7 pp	0.30	-12.2 pp	2.8 pp				

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 ending between October 1, 2021, and December 31, 2022 (PY 6 intervention).

Notes: The estimates in this exhibit are the result of a DDD model. DDD estimates that are significant at the 1%, 5%, or 10% significance levels are indicated by red, orange, or yellow shading, respectively. A full table of results, including sample counts and risk-adjusted means, is in **Appendix F: Claims-Based Health Equity Analyses, Exhibits F-13A and F-13B**, and outcome definitions are in **Appendix B: Data and Methods, Exhibit B-6**.

[a] The sample for this estimate failed tests for parallel trends in the baseline period. CJR = Comprehensive Care for Joint Replacement; DiD = difference-in-differences; DDD = difference-in-difference-in-differences; IRF = inpatient rehabilitation facility; LCI = lower confidence interval; pp = percentage point; PY = performance year; SNF = skilled nursing facility; UCI = upper confidence interval.

The CJR Model led to similar impacts on PAC utilization measures for patients who are Hispanic and patients who are Non-Hispanic White (**Exhibit V-20**). One notable difference was that patients who are Hispanic had a larger decrease in SNF length of stay of 1.7 days (p=0.34). However, the differential impact of –1.2 days (p=0.44) was not statistically significant.

Exhibit V-20: Patients Who Are Hispanic Had No Statistically Significant Impacts for Utilization Measures, nor Was Any Difference Statistically Significant Relative to Patients Who Are Non-Hispanic White

PAC utilization	Population	DiD (impact)	p-value (DiD)	90% LCI (DiD)	90% UCI (DiD)	DDD (differential impact)	p-value (DDD)	90% LCI (DDD)	90% UCI (DDD)
HH visits	Patients who are Hispanic	-0.7	0.51	-2.6	1.1	0.0	0.97	-1.6	1.5
	Patients who are Non-Hispanic White	-0.7[a]	0.23	-1.7	0.3				
SNF LOS	Patients who are Hispanic	-1.7	0.34	-4.6	1.3	-1.2	0.44	-3.8	1.4
	Patients who are Non-Hispanic White	-0.5	0.56	-1.9	0.9				
IRF LOS	Patients who are Hispanic	0.4	0.57	-0.8	1.7	0.0	0.95	-1.1	1.2
	Patients who are Non-Hispanic White	0.4	0.16	-0.1	0.8				

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 ending between October 1, 2021, and December 31, 2022 (PY 6 intervention).

Notes: The estimates in this exhibit are the result of a DDD model. DDD estimates that are significant at the 1%, 5%, or 10% significance levels are indicated by red, orange, or yellow shading, respectively. A full table of results, including sample counts and risk-adjusted means, is in **Appendix F: Claims-Based Health Equity Analyses, Exhibits F-14A and F-14B**, and outcome definitions are in **Appendix B: Data and Methods, Exhibit B-6**.

[a] The sample for this estimate failed tests for parallel trends in the baseline period. CJR = Comprehensive Care for Joint Replacement; DiD = difference-in-differences; DDD = difference-in-difference-in-differences; HH = home health; IRF = inpatient rehabilitation facility; LCI = lower confidence interval; pp = percentage point; SNF = skilled nursing facility; UCI = upper confidence interval.

The estimated impact on ED use for patients who are Hispanic was large and statistically significant: an increase of 3.1 percentage points ($p<0.01$) (**Exhibit V-21**). The differential impact of 3.0 percentage points relative to patients who are Non-Hispanic White was also statistically significant ($p<0.01$).

Patients who are Hispanic had a statistically insignificant estimated increase in the unplanned readmission rate, while patients who are Non-Hispanic White had a statistically significant decrease of 0.5 percentage points ($p=0.08$). The differential impact indicated that patients who are Hispanic were more likely to have unplanned readmissions by 1.9 percentage points ($p=0.03$) as a result of the CJR Model relative to patients who are Non-Hispanic White.

Exhibit V-21: Patients Who Are Hispanic Had a Significant Increase in ED Use, and This Was Significantly Different Relative to Patients Who Are Non-Hispanic White. They Also Had an Estimated Increase in Readmission Rate, Which Was Significantly Different Relative to Patients Who Are Non-Hispanic White

Measure	Population	DiD (impact)	p-value (DiD)	90% LCI (DiD)	90% UCI (DiD)	DDD (differential impact)	p-value (DDD)	90% LCI (DDD)	90% UCI (DDD)
Mortality rate	Patients who are Hispanic	0.3 pp	0.29	-0.1	0.7	0.3 pp	0.28	-0.1	0.7
	Patients who are Non-Hispanic White	0.0 pp	0.76	-0.1	0.1				
ED use	Patients who are Hispanic	3.1 pp	<0.01	1.7	4.5	3.0 pp	<0.01	1.5	4.5
	Patients who are Non-Hispanic White	0.1 pp	0.90	-0.7	0.8				
Readmission rate	Patients who are Hispanic	1.4 pp	0.11	0.0	2.8	1.9 pp	0.03	0.5	3.4
	Patients who are Non-Hispanic White	-0.5 pp	0.08	-1.0	0.0				

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 ending between October 1, 2021, and December 31, 2022 (PY 6 intervention).

Notes: The estimates in this exhibit are the result of a DDD model. DDD estimates that are significant at the 1%, 5%, or 10% significance levels are indicated by red, orange, or yellow shading, respectively. A full table of results, including sample counts and risk-adjusted means, is in **Appendix F: Claims-Based Health Equity Analyses, Exhibits F-12A and F-12B**, and outcome definitions are in **Appendix B: Data and Methods, Exhibit B-6**.

CJR = Comprehensive Care for Joint Replacement; DiD = difference-in-differences; DDD = difference-in-difference-in-differences; ED = emergency department; LCI = lower confidence interval; pp = percentage point; UCI = upper confidence interval.

Findings Across All Underserved Populations

Overall, our analysis of payments, utilization, and quality for underserved populations indicated that payments decreased for underserved populations and their reference populations. Although we did not estimate differential impacts for the components of total spending, we did observe a decrease in the intensity of the first PAC discharge destination, with limited evidence of decreases in SNF length of stay as well. For all four underserved populations, however, the underserved population became relatively more likely to be discharged to home *without* home health compared with their reference population, which is potentially concerning.

We also observed differential impacts in the quality metrics: Patients who are Black or African American and patients who are Hispanic had statistically significant increases in ED use in the 90 days after discharge. These impacts were significantly greater than those for patients who are Non-Hispanic White. In addition, patients who are Hispanic were relatively more likely to have unplanned readmissions than patients who are Non-Hispanic White due to the CJR Model.

c. Patient Mix Analysis for Underserved Populations

For the patient mix analysis, we analyzed how the underserved CJR patient populations have changed from baseline to intervention. We focused on the elective LEJR population, but we also conducted analyses for the all-LEJR population. The results tables for both populations studied, across all the underserved populations, can be found in **Appendix C: CJR Population Patient Characteristics, Section B.** Overall, we did not observe a clear pattern of change in patient characteristics for any of the underserved populations that we studied; however we highlight a few significant findings below.

Patients who are Black or African American were more likely to use home health or to have disability as the reason for Medicare coverage relative to patients who are Non-Hispanic White.[74] We found that patients who are Black or African American were less likely to have diabetes and more likely to have a Hierarchical Condition Category (HCC) count of 3 when compared with patients who are Non-Hispanic White. We did not observe a clear pattern of change in patient characteristics for the results. For patients who were dually eligible compared with patients who were not dually eligible, we again did not find any clear pattern of change based on age, demographics, HCC count, health status, or prior use. For patients who are Black or African American and dually eligible compared with those who are Non-Hispanic White and not dually eligible, we also found no consistent pattern of change in patient characteristics. For patients who are Hispanic compared with patients who are Non-Hispanic White, we identified positive differential impacts for two of the age bracket outcomes, patients who were 65 years or younger and patients who were 85 years and older.

E. Did the CJR Model Impact the Functional Status and Care Experiences for Underserved Populations in PY 6?

We examined changes in functional status and care experiences for LEJR patients in specific underserved populations using patient survey data from PY 6.

1. Summary of Findings

The callout box presents a summary of our findings. Results with an asterisk (*) are key findings.

- (*) We have found suggestive evidence (changes that are large but statistically insignificant) that the CJR Model has:
  - An adverse effect on functional outcomes for patients with dual eligibility
  - A beneficial effect on functional outcomes for patients who are Black or African American
  - A beneficial effect on satisfaction with care management and overall recovery for all underserved populations

[74] Although we did observe some statistically significant estimates, we are cautious about over-interpreting results with no clear pattern when conducting large numbers of statistical tests due to the multiple comparisons problem in statistical analysis.

- For both CJR and control patients, most measures showed large differences between underserved populations and the respective reference populations, with poorer outcomes for the underserved populations.
- The differences between underserved populations and the respective reference populations often showed most strongly in the highest (best) and lowest (worst) response options.
- There is little differential effect of the CJR Model on the need for caregiver help.

2. Methods

We studied three underserved populations: (1) patients who are dually eligible, compared with patients who are not dually eligible; (2) patients who are Black or African American, compared with patients who are Non-Hispanic White; and (3) patients who are Hispanic, compared with patients who are Non-Hispanic White. Our dataset contained 9,676 patient observations, including observations for 402 patients who were dually eligible for Medicaid, 370 patients who are Black or African American, and 307 patients who are Hispanic. Respondents self-identified their race and ethnicity, and many chose "prefer not to answer" to those questions: 10.0% for ethnicity and 9.6% for race. The sample counts were lower than might be expected due to this high rate of respondents choosing not to provide their self-identified race or ethnicity.

We analyzed responses to survey questions about functional status and pain, caregiver help, care management, and overall recovery.[75] We treated the ordered responses to these questions as ordinal data and used an ordinal logistic regression.

All regression specifications included indicators for CJR versus control group, for hip fracture, and for a variety of risk-adjusters: Medicaid eligibility, self-reported race and ethnicity, disability status, sex, age, self-reported income, HCC score, and presence of caregiver at home. In addition, all specifications controlled for the patient's functional status prior to surgery. We included interaction terms in the regressions to capture the effect of the CJR Model relative to the control group for each subpopulation and to capture the differential effect of CJR on an underserved population relative to the reference group.[76]

3. Results

The full set of results are presented in **Appendix E: Patient Survey Results**. They are summarized here.

[75] Full details on all the methods used for the analyses in this section can be found in **Appendix B: Data and Methods**.

[76] For the remainder of this section, we use the term *differences* to describe the effect of CJR on a subpopulation (difference between CJR and Control) and *differential effect* to describe the effects of the CJR Model on an underserved subpopulation relative to the effect of the CJR Model on a contrasting reference subpopulation.

a. Functional Status

The individual survey measures and results tables for functional status and pain are described in **Appendix B: Data and Methods and Appendix E: Patient Survey Results**.[77]

Both CJR and control patients, who were dually eligible had worse functional status outcomes than patients who were not dually eligible. All eight of the available functional status measures showed large adverse *differential* effects of the CJR Model for patients who were dually eligible: that is, CJR patient functional status declined by more than control patients. Two of the eight measures showed statistically significant adverse differential effects of the CJR Model: ability to walk by yourself without resting and difficulty getting on and off the toilet.

Both CJR and control patients who are Black or African American, had worse outcomes across all eight measures compared with patients who are Non-Hispanic White. However, the outcomes of patients who are Black or African American were *less worse* in CJR patients than in the control patient: suggesting CJR offered some beneficial or protective effect for Black or African American patients. For one of the eight measures, use of a mobility aid, the differential effect was statistically significant.

For both the CJR and control patients who are Hispanic, all measures showed large adverse differences (worse outcomes) in responses compared with patients who are Non-Hispanic White. There was no clear sign of differential effects of the CJR Model for patients who are Hispanic. One of the eight measures showed a statistically significant adverse differential effect of the CJR Model: medication use for pain in the joint you had replaced. Two other measures showed adverse differential effects, two showed beneficial differential effects, and one showed little effect.

b. Satisfaction With Care Management and Overall Recovery

Descriptions of the individual survey measures and results tables for care management and overall recovery are in **Appendix E: Patient Survey Results**. [78]

For both the CJR and control patients, all measures of satisfaction showed large adverse differences (worse outcomes) in responses between patients who were dually eligible and patients who were not dually eligible. Despite the adverse differential effects, we saw for measures of functional status, all five measures of satisfaction showed beneficial *differential* effects of the CJR Model for patients who were dually eligible, although none of the five measures showed statistically significant differential effects of the CJR Model. That is, the outcomes of patients who were dually eligible were *less worse* in CJR patients than in control patients.

For the control patients who are Black or African American, all measures showed large adverse differences (worse outcomes) in responses compared with control patients who are Non-Hispanic White. This pattern was not as clear for the CJR patients. All five measures showed beneficial

[77] **Appendix B: Data and Methods, Exhibit B-13** provides detail on the survey measures related to functional status and pain. The results of our analyses for functional status are presented in **Exhibit E-17**, **Exhibit E-18**, **and Exhibit E-19**.

[78] **Appendix B: Data and Methods, Exhibit B-13** provides detail on the survey measures related to satisfaction with care management and overall recovery. The results of our analyses for satisfaction with care management and overall recovery are presented in **Exhibit E-20**, **Exhibit E-21**, **and Exhibit E-22**.

differential effects of the CJR Model for patients who are Black or African American, but none of the five measures showed statistically significant differential effects.

For the control patients who are Hispanic, all measures showed large adverse differences (worse outcomes) in responses compared with control patients who are Non-Hispanic White. This pattern was not as evident for the CJR patients. All five measures showed beneficial *differential* effects of the CJR Model for patients who are Hispanic. This was more than we would expect from chance, but again, none of these results were statistically significant.

c. Need for Caregiver Help

The individual survey measure results tables for caregiver help are in **Appendix E: Patient Survey Results**.[79]

For both the CJR and control patients, all measures of caregiver help showed large adverse differences (worse outcomes) in responses compared with patients who were dually eligible and patients who were not dually eligible. All three measures showed adverse *differential* effects of the CJR Model for patients who were dually eligible, but none of the measures showed statistically significant differential effects. This result may correspond to the adverse differential effects on measures of functional status and pain for patients who were dually eligible.

For both the CJR and control patients who are Black or African American, all measures showed large adverse differences (worse outcomes) in responses compared with patients who are Non-Hispanic White. There was no clear sign of a differential effect of the CJR Model for patients who are Black or African American.

For both the CJR and control patients who are Hispanic, all measures showed large adverse differences (worse outcomes) in responses compared with patients who are Non-Hispanic White, but again, there was no clear pattern of a differential effect of the CJR Model.

[79] **Appendix B: Data and Methods, Exhibit B-13** provides detail on the survey measures related to caregiver help. The results of our analyses for caregiver help are presented in **Exhibit E-23**, **Exhibit E-24**, **and Exhibit E-25**.

VI. Analysis of Potential Unintended Consequences of the CJR Model

In this chapter, we present analyses that investigated potential unintended consequences of the Comprehensive Care for Joint Replacement (CJR) Model. We examined whether characteristics of the patient population changed by comparing demographics, health status, and prior use for CJR and control patients at baseline and in Performance Year 6 (PY 6). The updated target price methodology in PY 6 included additional risk adjustment based on age, dual-eligibility status, and count of Hierarchical Condition Categories (HCCs). For this analyses, we examined changes in patient characteristics that could be associated with risk and selection. To investigate potentially delayed care, we looked at 30-day post-episode payments. To understand whether the model could be increasing or reducing the number of lower extremity joint replacements (LEJRs) that hospitals performed, we studied changes in LEJR rates in CJR and control metropolitan statistical areas (MSAs).

Acronyms	
ACH	acute care hospital
CI	confidence interval
CJR	Comprehensive Care for Joint Replacement
DiD	difference-in-differences
ESRD	end-stage renal disease
FFS	fee-for-service
HCC	Hierarchical Condition Category
HH	home health
IRF	inpatient rehabilitation facility
LCI	lower confidence interval
LEJR	lower extremity joint replacements
MSA	metropolitan statistical area
MS-DRG	Medicare Severity-Diagnosis Related Group
pp	percentage point
PY	performance year
SNF	skilled nursing facility
UCI	upper confidence interval

A. Did the CJR Patient Population in the PY 6 Intervention Period Differ from the Baseline Period?

We analyzed whether the CJR patient population changed from baseline to PY 6 relative to the control patient population in a manner that would reflect unintended participant behavior. The goal was to identify whether changes in the patient populations for CJR hospitals and control hospitals were comparable over time. As discussed in **Chapter I: CJR Model Background,** the CJR episode target price was updated in 2021.[80] Starting in PY 6, CMS adjusted CJR target prices to account for additional patient characteristics such as age, dual-eligibility status, and HCC counts. This policy change may have incentivized hospitals to select patients depending on which patient characteristics would improve their chances of earning reconciliation, if the hospital believed patients with certain characteristics were more likely to have average episode payments below or above its corresponding target price. We examined characteristics directly related to the target pricing as well as other characteristics related to demographics, health status, and prior care.

[80] See the CJR Three-Year Extension Final Rule for details on target pricing changes: https://www.federalregister.gov/documents/2021/05/03/2021-09097/medicare-program-comprehensive-care-for-joint-replacement-model-three-year-extension-and-changes-to

1. Summary of Findings

- We found modest evidence that patient complexity increased for both CJR and control elective patient populations from the baseline period to PY 6.
- We did not observe notable changes in the CJR hip fracture patient population relative to the changes in the control hip fracture patient population.
- We also did not find evidence that the CJR elective patient population changed from PY 5 to PY 6 relative to the control patient population for patient characteristics adjusted for in the PY 6 target pricing.

2. Methods

We examined changes in the characteristics of CJR patients from the baseline to PY 6 relative to control patients. Using unadjusted difference-in-differences (DiD) models, we compared the mean across time (baseline to PY 6) and group (CJR to control) for each patient characteristic. See **Appendix B: Data and Methods** for additional information.

We assessed changes in demographics, Medicare eligibility, HCCs (both average score and count), health status, and health care use prior to the LEJR procedure. For each characteristic, exhibits show baseline and PY 6 averages for the CJR and control populations and the calculated change in the CJR–control difference between the baseline and PY 6. We conducted analyses for two population groups: patients receiving elective LEJRs and patients receiving LEJRs due to a hip fracture. For this analysis, we defined elective LEJRs as those with Medicare Severity-Diagnosis Related Group (MS-DRG) 470 or in the outpatient setting.[81] This excluded MS-DRG 469 episodes.[82]

To better isolate changes in patient populations that may have occurred in response to the changes to the target price methodology, we conducted an additional analysis examining changes in patients between PY 5 and PY 6 for elective MS-DRG 470 and outpatient LEJRs. For this analysis, we estimated unadjusted DiD models on patient characteristics using PY 5 as the "baseline period" and PY 6 as the "intervention period."

3. Results

In our first analysis, we examined the relative change in patient characteristics for elective MS-DRG 470 and outpatient patients for the CJR and control populations. We were interested in studying how the CJR patient population changed from baseline to PY 6 relative to the control group. We focused on elective MS-DRG 470 and outpatient patients because they had the most common type of LEJR procedure and were generally less complex than other LEJR patients (MS-DRG 469 or hip fracture). Presurgery patient optimization and related activities by CJR hospitals

[81] MS-DRG 470 is defined as major joint replacement or reattachment of lower extremity *without* major complications and comorbidities and excludes all procedures with a principal diagnosis of hip fracture.

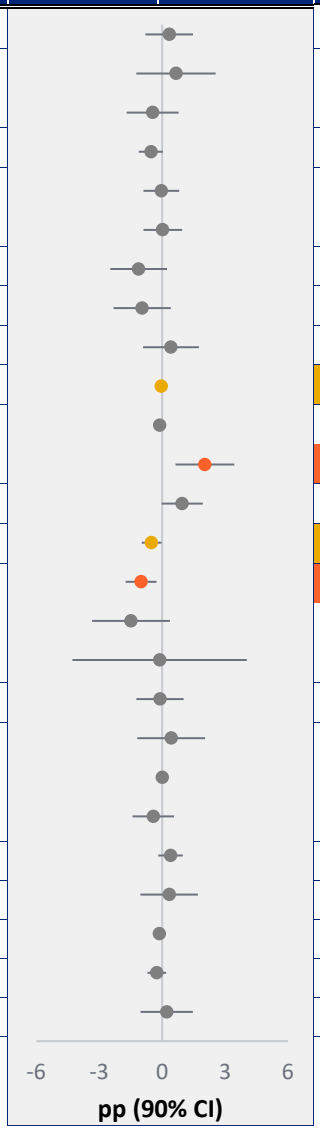
[82] MS-DRG 469 is defined as major joint replacement or reattachment of lower extremity *with* major complications and comorbidities.

may have resulted in CJR hospitals having a less complex patient population relative to control hospitals.

As shown in **Exhibit VI-1**, we found modest evidence that both CJR and control hospitals had a relative shift toward patients with higher complexity, as measured by our HCC score characteristics. However, the increase was less for CJR hospitals, so that CJR hospital patients became relatively less risky. For the control group, we saw a slightly larger increase in the proportion of patients with three or more HCCs, relative to the CJR group. Average HCC scores decreased by 0.04 more points for CJR hospitals than for control hospitals from the baseline (p=0.10). We observed the same pattern for the proportion of patients with varying numbers of HCCs: a relative *increase* of 2.0 percentage points (p=0.02) with *zero* HCCs (patients with no chronic conditions) and a relative *decrease* in counts of two, three, and four or more HCCs (–0.5 percentage points, p=0.08; –1.0 percentage points, p=0.02; –1.5 percentage points, p=0.19, respectively). There were no notable relative changes in other characteristics for the CJR patient population, including health status and prior health care usage.

Exhibit VI-1: CJR Elective MS-DRG 470 and Outpatient Patient Population Became Less Complex Between Baseline and PY 6 Relative to Control Patients, as Measured by HCC Variables

Characteristic		Proportion				Change from baseline to PY 6 for CJR vs. change from baseline to PY 6 for control			
		CJR		Control		Estimate (pp)	Less frequent	More frequent	p-value
		Base	PY 6	Base	PY 6				
Age	<65 years	8.5%	3.8%	8.8%	3.7%	0.3			0.63
	65–74 years	49.3%	53.1%	51.9%	55.0%	0.7			0.57
	75–84 years	34.8%	37.2%	32.9%	35.7%	-0.5			0.55
	85+ years	7.4%	6.0%	6.5%	5.5%	-0.5			0.12
Sex	Female	64.7%	62.7%	64.4%	62.5%	-0.0			0.94
Race and ethnicity	Black or African American	6.4%	4.2%	7.4%	5.2%	0.0			0.95
	Hispanic	6.5%	5.1%	2.9%	2.7%	-1.1			0.17
Eligibility	Eligible for Medicaid	12.3%	6.5%	9.8%	4.9%	-1.0			0.25
	Disability, no ESRD	16.1%	10.8%	16.4%	10.7%	0.4			0.60
HCC	Score[a]	1.25	1.32	1.17	1.29	-0.04			0.10
	Count[b]	2.27	2.51	2.10	2.45	-0.11			0.12
	Count: 0	19.1%	17.4%	21.6%	17.9%	2.0			0.02
	Count: 1	24.3%	22.3%	25.8%	22.8%	1.0			0.11
	Count: 2	20.5%	19.5%	20.0%	19.5%	-0.5			0.08
	Count: 3	14.2%	14.6%	13.2%	14.6%	-1.0			0.02
	Count: 4+	22.0%	26.3%	19.5%	25.3%	-1.5			0.19
Health status	Obesity	17.3%	36.4%	18.0%	37.2%	-0.1			0.96
	Diabetes	29.3%	26.6%	27.1%	24.6%	-0.1			0.88
	Hypertension	75.1%	76.4%	75.2%	76.1%	0.4			0.66
	Dementia	3.2%	2.5%	3.1%	2.5%	0.0			0.97
	Congestive heart failure	12.3%	12.5%	11.4%	12.0%	-0.4			0.48
Prior use	ACH stay	11.0%	5.8%	11.1%	5.5%	0.4			0.26
	HH use	10.5%	8.1%	9.8%	7.1%	0.3			0.69
	IRF stay	1.1%	0.5%	1.1%	0.6%	-0.1			0.43
	SNF stay	3.6%	1.3%	3.2%	1.1%	-0.3			0.35
	Any prior care	26.2%	20.9%	25.9%	20.3%	0.2			0.78



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 ending between October 1, 2021, and December 31, 2022 (PY 6 intervention).

Notes: The estimates in this exhibit are the results of unadjusted DiD models on patient characteristics. Estimates that are significant at the 1%, 5%, or 10% significance level are indicated by red, orange, or yellow shaded cells, respectively. Count characteristics refer to the count of individual HCCs. Any prior care included inpatient hospital, psychiatric hospital, emergency department, SNF, IRF, HH, long-term care hospital, and hospice use during the 6 months prior to the anchor hospitalization. ACH = acute care hospital; CI = confidence interval; CJR = Comprehensive Care for Joint Replacement; DiD = difference-in-differences; ESRD = end-stage renal disease; HCC = Hierarchical Condition Category; HH = home health; IRF = inpatient rehabilitation facility; MS-DRG = Medicare Severity-Diagnosis Related Group; pp = percentage point; PY = performance year; SNF = skilled nursing facility.

[a] Estimates for HCC score are expressed as units of the score instead of percentage points.

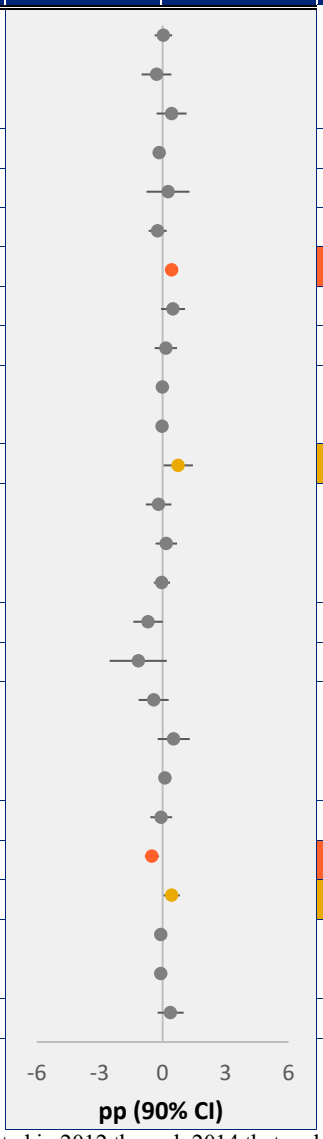
[b] Estimates for the count of HCCs are expressed as numbers instead of percentage points.

Our second analysis also focused on elective MS-DRG 470 and outpatient patients but compared patient characteristics of LEJR patients from PY 5 to PY 6. Starting in PY 6, CMS risk-adjusted target prices for patient age, dual-eligibility status, and counts of HCCs. We performed this analysis to study any immediate observable changes that may have resulted from this adoption of patient-level risk adjustment in the target prices.

The results shown in **Exhibit VI-2** indicate no clear pattern of change in patient characteristics. Although some individual characteristics had relative changes that were statistically significant, the changes were generally inconsistent with other characteristics. For example, we saw a relative 0.7 percentage point increase (p=0.08) in patients with zero HCCs but also observed a relative decrease in HCC score (-0.01 points, p=0.34) and decreases in many characteristics related to a patient's health status. Similarly, for prior use, we saw a relative increase in home health use (0.4 percentage points, p=0.07) but a relative decrease in any inpatient hospital stay (-0.5 percentage points, p=0.01). Notably, changes in the proportion of patients in different age groups and the proportion of patients dually eligible for Medicaid and Medicare were similar for the CJR and control populations.

Exhibit VI-2: Between PY 5 and PY 6, the Elective MS-DRG 470 and Outpatient Patient Population Had Similar Changes for CJR and Control Hospitals

Characteristic		Proportion				Change from PY 5 to PY 6 for CJR vs. change from PY 5 to PY 6 for control			
		CJR		Control		Estimate (pp)	Less frequent	More frequent	p-value
		PY 5	PY 6	PY 5	PY 6				
Age	<65 years	4.9%	3.8%	4.9%	3.7%	0.0			0.88
	65–74 years	56.2%	53.1%	57.8%	55.0%	-0.3			0.50
	75–84 years	33.3%	37.2%	32.3%	35.7%	0.4			0.33
	85+ years	5.6%	6.0%	5.0%	5.5%	-0.2			0.29
Sex	Female	61.3%	62.7%	61.3%	62.5%	0.3			0.68
Race and ethnicity	Black or African American	4.6%	4.2%	5.4%	5.2%	-0.2			0.37
	Hispanic	4.5%	5.1%	2.5%	2.7%	0.4			0.02
Eligibility	Eligible for Medicaid	6.8%	6.5%	5.7%	4.9%	0.5			0.16
	Disability, no ESRD	12.2%	10.8%	12.2%	10.7%	0.2			0.63
HCC	Score[a]	1.30	1.32	1.26	1.29	-0.0			0.34
	Count[b]	2.46	2.51	2.36	2.45	-0.0			0.24
	Count: 0	18.2%	17.4%	19.4%	17.9%	0.7			0.08
	Count: 1	22.9%	22.3%	23.2%	22.8%	-0.2			0.61
	Count: 2	19.2%	19.5%	19.4%	19.5%	0.2			0.57
	Count: 3	14.1%	14.6%	14.1%	14.6%	-0.0			0.89
	Count: 4+	25.6%	26.3%	23.9%	25.3%	-0.7			0.10
Health status	Obesity	36.3%	36.4%	36.0%	37.2%	-1.2			0.16
	Diabetes	26.7%	26.6%	24.2%	24.6%	-0.4			0.32
	Hypertension	74.9%	76.4%	75.2%	76.1%	0.5			0.26
	Dementia	2.3%	2.5%	2.4%	2.5%	0.1			0.49
	Congestive heart failure	11.7%	12.5%	11.2%	12.0%	-0.1			0.83
Prior use	ACH stay	7.7%	5.8%	6.9%	5.5%	-0.5			0.01
	HH use	8.3%	8.1%	7.7%	7.1%	0.4			0.07
	IRF stay	0.5%	0.5%	0.6%	0.6%	-0.1			0.19
	SNF stay	1.5%	1.3%	1.2%	1.1%	-0.1			0.46
	Any prior care	21.2%	20.9%	21.0%	20.3%	0.4			0.31



Source: CJR evaluation team analysis of Medicare claims and enrollment data for episodes initiated in 2012 through 2014 that ended between January 1, 2020, and September 30, 2021 (PY 5 intervention) and episodes ending between October 1, 2021, and December 31, 2022 (PY 6 intervention).

Notes: The estimates in this exhibit are the results of unadjusted DiD models on patient characteristics using PY 5 as the "baseline period" and PY 6 as the "intervention period." Estimates that are significant at the 1%, 5%, or 10% significance level are indicated by red, orange, or yellow shaded cells, respectively. Count characteristics refer to the count of individual HCCs. Any prior care included inpatient hospital, psychiatric hospital, emergency department, SNF, IRF, HH, long-term care hospital, and hospice use during the 6 months prior to the anchor hospitalization. ACH = acute care hospital; CI = confidence interval; CJR = Comprehensive Care for Joint Replacement; DiD = difference-in-differences; ESRD = end-stage renal disease; HCC = Hierarchical Condition Category; HH = home health; IRF = inpatient rehabilitation facility; MS-DRG = Medicare Severity-Diagnosis Related Group; pp = percentage point; PY = performance year; SNF = skilled nursing facility.

[a] Estimates for HCC score are expressed as units of the score instead of percentage points.

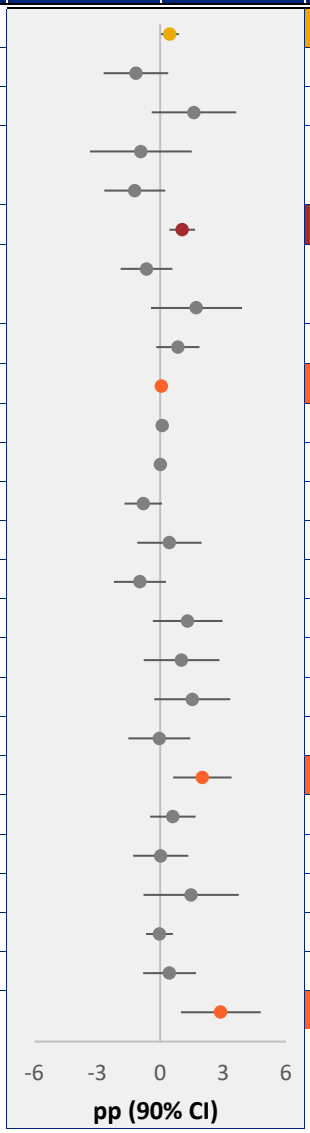
[b] Estimates for the count of HCCs are expressed as numbers instead of percentage points.

In our third analysis, we compared changes in the characteristics of fracture patients in the baseline period versus PY 6 for CJR hospitals compared with control hospitals. Hip fracture LEJRs present hospitals with limited opportunities for presurgery patient optimization or discretion in scheduling of the surgery; thus, patient characteristics may have little influence on hospitals' patient selection for this population. Relative changes in the CJR hip fracture patient population likely represent general changes or fluctuations in the hospitals' patient populations occurring beyond the CJR Model.

As shown in **Exhibit VI-3**, we found no clear pattern of change in patient characteristics for the hip fracture population, but there were a few notable changes. CJR hospitals saw a relative increase in the proportion of hip fracture patients who are Black or African American (1.1 percentage points, $p<0.01$) from the baseline to PY 6. In addition, changes in certain characteristics indicated that CJR hospitals may have had a more complex hip fracture patient population relative to control hospitals: a relative increase of 0.07 ($p=0.05$) in HCC scores, a relative increase of 2.0 percentage points ($p=0.02$) in the proportion of patients with dementia, and a relative increase of 2.9 percentage points ($p=0.01$) in the number of patients using any type of health care prior to their hip fracture. Other patient characteristics related to health status, prior use, and HCC variables had only small relative changes or changes that were not statistically significant.

Exhibit VI-3: Changes in the CJR Hip Fracture Patient Population Between Baseline and PY 6 Were Similar to Changes at Control Hospitals

Characteristic		Proportion				Change from baseline to PY 6 for CJR vs. change from baseline to PY 6 for control			
		CJR		Control		Estimate (pp)	Less frequent	More frequent	p-value
		Base	PY 6	Base	PY 6				
Age	<65 years	3.0%	2.2%	3.4%	2.1%	0.5			0.07
	65–74 years	15.6%	18.0%	16.4%	20.0%	-1.2			0.22
	75–84 years	31.4%	35.0%	33.3%	35.3%	1.6			0.19
	85+ years	50.1%	44.9%	46.8%	42.6%	-0.9			0.53
Sex	Female	73.9%	70.2%	73.5%	70.9%	-1.2			0.17
Race and ethnicity	Black or African American	3.2%	3.0%	4.7%	3.4%	1.1			<0.01
	Hispanic	5.8%	5.1%	2.2%	2.2%	-0.7			0.39
Eligibility	Eligible for Medicaid	19.8%	16.9%	16.1%	11.4%	1.7			0.19
	Disability, no ESRD	9.8%	10.4%	10.8%	10.6%	0.8			0.18
HCC	Score[a]	2.42	2.60	2.38	2.49	0.1			0.05
	Count[b]	4.60	4.98	4.54	4.82	0.1			0.14
	Count: 0	0.3%	0.3%	0.2%	0.2%	0.0			0.86
	Count: 1	9.6%	9.3%	9.9%	10.5%	-0.8			0.14
	Count: 2	15.5%	13.4%	15.9%	13.4%	0.4			0.64
	Count: 3	16.7%	14.5%	16.9%	15.6%	-1.0			0.20
	Count: 4+	58.0%	62.6%	57.1%	60.4%	1.3			0.19
Health status	Obesity	4.2%	10.8%	5.2%	10.8%	1.0			0.35
	Diabetes	29.3%	30.1%	26.9%	26.2%	1.5			0.16
	Hypertension	75.2%	76.5%	75.3%	76.6%	-0.0			0.96
	Dementia	29.7%	27.2%	31.4%	26.8%	2.0			0.02
	Congestive heart failure	26.0%	23.9%	25.2%	22.4%	0.6			0.36
Prior use	ACH stay	22.2%	19.0%	22.4%	19.2%	0.0			0.98
	HH use	25.9%	22.5%	26.1%	21.2%	1.5			0.28
	IRF stay	2.7%	2.5%	3.4%	3.2%	-0.0			0.93
	SNF stay	12.4%	11.0%	11.8%	9.9%	0.4			0.56
	Any prior care	48.0%	46.1%	49.1%	44.3%	2.9			0.01



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 ending between October 1, 2021, and December 31, 2022 (PY 6 intervention).

Notes: The estimates in this exhibit are the result of unadjusted DiD models on patient characteristics. Estimates that are significant at the 1%, 5%, or 10% significance level are indicated by red, orange, or yellow shaded cells, respectively. Count characteristics refer to the count of individual HCCs. Any prior care included inpatient hospital, psychiatric hospital, emergency department, SNF, IRF, HH, long-term care hospital, and hospice use during the 6 months prior to the anchor hospitalization or procedure. ACH = acute care hospital; CI = confidence interval; CJR = Comprehensive Care for Joint Replacement; DiD = difference-in-differences; ESRD = end-stage renal disease; HCC = Hierarchical Condition Category; HH = home health; IRF = inpatient rehabilitation facility; pp = percentage point; PY = performance year; SNF = skilled nursing facility.

[a] Estimates for HCC score are expressed as units of the score instead of percentage points.

[b] Estimates for the count of HCCs are expressed as numbers instead of percentage points.

Results from our additional patient mix analyses of elective LEJR patients for health equity subpopulations are in **Appendix C: CJR Population Patient Characteristics, Section B**.

B. Did the CJR Model Impact the Payments in the 30 Days Following the Episode?

Under the CJR Model, CJR hospitals were financially accountable for the cost and quality of health care services during a 90-day episode of care following an LEJR procedure. The episode of care began with the hospitalization for the surgery and extended through the 90 days after hospital discharge (including the date of discharge). Any services provided immediately after the 90 days were not included in the episode; thus, there could have been a financial incentive for CJR hospitals to delay care until the end of the episode. To study this possible unintended consequence, we monitored payments for services provided up to 30 days *after* the episode to identify whether CJR hospitals postponed services to reduce episode payments. Postponing services could have implications for patients' long-term health. Moreover, because the outcomes we used to study the impact of the CJR Model were created for the episode period, if CJR hospitals postponed services, we could have overestimated the reductions in payments and savings to Medicare.

1. Summary of Findings

- For hip fracture patients, CJR hospitals increased Medicare spending 30 days after the episode by $351 per episode relative to control hospitals.

2. Methods

For this analysis, we used the same DiD design we employed to estimate the impact of the model on payments, utilization, and quality of care. We estimated the change in post-episode payments for LEJR episodes between the baseline (April 2012 through March 2015) and PY 6 (October 2021 through December 2022) initiated at mandatory CJR hospitals relative to those initiated at control group hospitals. Our measure of post-episode payments included all Medicare standardized allowed amounts for all health care services covered under Medicare Parts A and B during the 30-day period immediately following the LEJR episode, 90 to 120 days after discharge. We risk-adjusted estimates for patient, market, and hospital characteristics. We weighted the control group MSAs to be representative of the distribution of the mandatory CJR MSAs. We provide a full description of the methods used in **Appendix B: Data and Methods**.

3. Results

We estimated the impact of the CJR Model during PY 6 on post-episode payments for three populations: all LEJR episodes, elective LEJR episodes, and hip fracture episodes. If CJR hospitals were intentionally postponing services to meet their target prices, we would expect to find a relative increase in payments after the episode period. For all LEJR episodes, we found that the

CJR Model led to a relative increase in payments 30 days after the episode of $72 (p=0.04) **(Exhibit VI-4)**.[83]

Exhibit VI-4: The CJR Model Led to an Increase in 30-Day Post-Episode Payments for Hip Fracture LEJRs

Population	Impact	Impact as a percentage[a]	p-value	90% LCI	90% UCI
All LEJR episodes[b]	$72	5.0%	0.04	$14	$129
Elective LEJR episodes[b]	$21	1.8%	0.53	-$34	$75
Hip fracture LEJR episodes	$351	10.4%	0.05	$51	$650

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 ending between October 1, 2021, and December 31, 2022 (PY 6 intervention).

Notes: The estimates in this exhibit were the result of a DiD model. DiD estimates that were significant at the 1%, 5%, or 10% significance level are indicated by red, orange, or yellow shading, respectively. A full table of results, including sample counts and risk-adjusted means, is in **Appendix D: Claims-Based Impact Analyses – Detailed Tables, Exhibit D-15.** CJR = Comprehensive Care for Joint Replacement; DiD = difference-in-differences; LCI = lower confidence interval; LEJR = lower extremely joint replacement; PY = performance year; UCI = upper confidence interval.

[a] Percentages were calculated by dividing the impact by the risk-adjusted mean value of the outcome for CJR episodes in the baseline period.

[b] Indicates that we believe the CJR and control populations may have been on relatively large differential trends in the baseline period for this outcome

When examining changes in post-episode payments by episode type, we found that hip fracture episodes were driving the increase in post-episode payments in the all-LEJR population. For elective episodes, CJR had no impact on post-episode spending (p=0.53), but for hip fracture episodes, CJR had a relative increase in payments 30 days after the episode (p=0.05). Although this relative increase of $351 was small compared with the average fracture episode total payment,[84] it was an approximate 10% increase in 30-day post-episode payments from the baseline.

Given that we only observed this result for patients with hip fractures, the potential implications were unclear. Hip fracture patients were older, more likely to have comorbidities, and more likely to have used health care services prior to their LEJR surgery than elective patients and had a 25% mortality rate 1 year after their LEJR surgery.[85] We cannot assess if the increase reflects intentional delays in care or longer, potentially appropriate, recovery periods. Survey and claims-based quality data do not detect any differences in patient experience: we found that hip fracture patients at CJR hospitals reported similar quality of care to hip fracture patients at control hospitals during the episode. CJR hip fracture patients showed improvements in certain measures of

[83] There was evidence that the CJR and control groups were on differential trends in the baseline period for this outcome, so we do not believe our DiD estimate of $72 was an unbiased causal estimate of the CJR Model. However, the differential pre-trends were in the direction that had they continued absent the CJR Model, our impact would have been an underestimate of the relative increase in post-episode spending. See **Appendix D: Claims-Based Impact Analyses – Detailed Tables** for more discussion on our assessment of parallel trends.

[84] The average risk-adjusted hip fracture LEJR episode was $48,257 in PY 6 for CJR hospitals.

[85] Centers for Medicare & Medicaid Services. (2023). *Comprehensive Care for Joint Replacement Model - Fifth Annual Report.* https://www.cms.gov/priorities/innovation/data-and-reports/2023/cjr-py5-annual-report.

functional status shortly after the episode period relative to control hip fracture patients.[86] We will continue to monitor these results.

C. Did the CJR Model Impact the Total Market Volume of Elective LEJR Discharges for Beneficiaries Residing in CJR MSAs?

Over the past few decades, Americans have been receiving more LEJR surgeries.[87,88,89] Given this growth, one concern is that the CJR Model itself could boost LEJR volume beyond what it would have been absent the model by making the surgery more financially rewarding to CJR hospitals. If CJR hospitals reduced their average episode payments by providing elective LEJRs to patients who otherwise would have foregone or delayed the procedure, Medicare savings due to the CJR Model would be offset by the payments for these additional LEJR surgeries. In this section, we discuss how LEJR surgery rates have changed for the Medicare fee-for-service (FFS) population in CJR MSAs compared with control MSAs.

1. Summary of Findings

- For beneficiaries living in CJR MSAs, there was a relative increase in the number of LEJRs during PY 6 of 47 LEJRs per 100,000 Medicare FFS beneficiaries.

2. Methods

This analysis estimated the impact of the CJR Model on the probability of receiving at least one elective LEJR for *all* Medicare FFS beneficiaries living in mandatory CJR MSAs versus control MSAs in 2012 to 2014 (the baseline of this analysis) and 2022 (spanning most of PY 6). This method was the same as that used for the analyses presented in **Chapter V: Health Equity Impact of the CJR Model**, but we analyzed the FFS population as a whole and did not compare impacts across the subpopulations studied in the health equity chapter.

In this analysis, we included *all* elective LEJRs, regardless of the location of the procedure and whether that hospital was in the CJR or control group. LEJRs included those that took place in the inpatient and outpatient setting as well as those performed in ambulatory surgical centers. We again considered only elective LEJRs because hip fracture LEJRs present hospitals with limited opportunities for presurgery patient optimization or discretion in scheduling of the surgery. We report results in terms of the probability of a Medicare FFS beneficiary receiving at least one elective LEJR in a given year, which we referred to and interpreted as LEJR rates per 100,000 FFS beneficiaries per year.

---

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

87 Wolford, M. L., Palso, K., & Bercovitz, A. (2015). *Hospitalization for total hip replacement among inpatients aged 45 and over: United States, 2000–2010*. NCHS data brief, no. 186. National Center for Health Statistics. https://www.cdc.gov/nchs/data/databriefs/db186.pdf

88 Williams, S. N., Wolford, M. L., & Bercovitz, A. (2015). *Hospitalization for total knee replacement among inpatients aged 45 and over: United States, 2000–2010*. NCHS data brief, no. 210. National Center for Health Statistics. https://www.cdc.gov/nchs/data/databriefs/db210.pdf

89 Centers for Medicare & Medicaid Services. (2021). *Comprehensive Care for Joint Replacement Model - Fourth Annual Report*. https://www.cms.gov/priorities/innovation/data-and-reports/2021/cjr-py4-annual-report

3. Results

For beneficiaries living in mandatory CJR MSAs, we found that the CJR Model had a small relative increase in the rate of elective LEJRs of 47 LEJRs per 100,000 beneficiaries (p=0.10) (**Exhibit VI-5**). To put this in context, in 2022, there were 1,430 LEJRs per 100,000 beneficiaries in these MSAs. Although the increase we identified was statistically significant at the 10% level, even the upper bound of the 90% confidence interval was small, at 94 LEJRs per 100,000 beneficiaries.

Exhibit VI-5: For Patients Residing in CJR MSAs, the CJR Model Had a Small Increase in Elective LEJR Volume

Time period	Impact on volume (LEJRs per 100,000 beneficiaries)	p-value	90% LCI	90% UCI
Calendar Year 2022 (most of PY 6)	47	0.10	0	94

Source: CJR evaluation team analysis of Medicare claims and enrollment data for Medicare FFS beneficiaries enrolled in Medicare between 2012 and 2015 (baseline) or during 2022 (intervention).

Notes: The estimates in this exhibit were the result of a DiD model on the probability of receiving at least one elective LEJR for all Medicare FFS beneficiaries living in mandatory CJR MSAs or their corresponding control MSAs. Estimates were reported as number of LEJRs per 100,000 Medicare FFS beneficiaries. Estimates that were significant at the 1%, 5%, or 10% significance level are indicated by red, orange, or yellow shaded cells, respectively. CJR = Comprehensive Care for Joint Replacement; LCI = lower confidence interval; LEJR = lower extremity joint replacement; PY = performance year; UCI = upper confidence interval.

If the financial incentives of the CJR Model caused CJR hospitals to perform more LEJRs, we would expect hospitals to provide elective LEJRs to lower-cost beneficiaries who otherwise would have foregone or delayed the procedure, reducing their average payments relative to target prices and earning additional reconciliation payments. The results we found for PY 6 support this hypothesis, and although the increase in elective LEJR volume was small in magnitude, it potentially could have helped CJR hospitals earn additional reconciliation payments. Starting in PY 6, however, the CJR Model used a target pricing methodology that accounted for several patient characteristics, which may diminish the financial reward from increasing the number of elective LERJs to lower-cost beneficiaries. What may be causing this small relative increase in the elective LEJR rate is unclear.

VII. Accountable Care Organization (ACO) Experiences of CJR Hospitals

In Performance Year 6 (PY 6), the qualitative research team conducted telephone interviews with mandatory Comprehensive Care for Joint Replacement (CJR) participant hospitals to explore how the CJR Model hospitals are affected by their experiences working with Accountable Care Organizations (ACOs). This effort aimed to expand our understanding of how value-based care models and initiatives work together and provide insight for Centers for Medicare & Medicaid Services (CMS) on how they can be improved.

Acronyms	
ACO	Accountable Care Organization
AHA	American Hospital Association
BPCI	Bundled Payments for Care Improvement
CJR	Comprehensive Care for Joint Replacement
CMS	Center for Medicare & Medicaid Services
LEJR	lower extremity joint replacement
PAC	post-acute care
PY	performance year
REACH	Realizing Equity, Access, and Community Health
SDOH	social determinants of health

A. Methods

1. Protocols

We developed and implemented a 45-minute semi-structured interview guide that included questions tailored to the hospital's reported experience with Medicare ACOs. If the hospital did not participate in a Medicare ACO, interviewees were asked about ACOs more broadly, including any experiences with commercial ACOs or treating patients who are part of another ACO. The key questions were:

- What does it mean for a CJR hospital to be a part of a Medicare ACO?
- How are CJR hospitals influenced by their participation in a Medicare ACO?
- Are actions taken to respond to Medicare ACOs similar to or different from those taken to respond to the CJR Model?
- What are the benefits and challenges of participating in a Medicare ACO and the CJR Model concurrently?
- What are the experiences with ACOs for CJR hospitals that do not participate in a Medicare ACO?

Prior to the interview, hospital representatives were asked to complete a brief web-based survey to gather information about the hospital's experience with ACOs, including any participation in a Medicare ACO. We used responses from the preinterview surveys to tailor the interview protocols.

2. Interviewee Selection and Recruitment

The team interviewed representatives from current CJR participant hospitals (as of September 2023). We used 2021 American Hospital Association (AHA) Survey data to randomly sample 50 CJR hospitals that reported leading a Medicare ACO, 25 CJR hospitals that participate but do not lead a Medicare ACO or that lead a non-Medicare ACO, and 25 CJR hospitals that do not participate in an ACO.

The unique Medicare identification numbers of the 100 sampled hospitals were linked to an updated CJR participant list obtained from CMS that included point of contact name, email address, and telephone number. We refined our outreach approach based on our prior rounds of telephone interviews. We sent an initial email invitation to each hospital in our sample, asking whether they would participate in this round of interviews. We included a description of the topic of interest and attached a frequently asked questions document and informed consent information to the email. Once the hospital responded, we followed up with a confirmation email and requested that the hospital point of contact complete the preinterview survey to inform and tailor the interview.

Outreach was conducted in three waves to monitor our response rate and adjust outreach strategy if needed, for example recruiting additional ACO participants if needed. Across our first two waves, 65 hospitals were invited to our interviews. Hospital response rate was higher than anticipated during the first two waves of randomly selected hospitals from our sample, so the last wave was a smaller (8 hospital) purposively sampled selection of AHA-surveyed Medicare ACO participant hospitals to ensure we had adequate representation from that group of hospitals. Our team conducted outreach to 73 hospitals from the 100 sampled, and 32 hospitals agreed to participate in an interview.

3. Data Collection

We conducted interviews with 32 hospitals between September and November 2023 (**Exhibit VII-1**). Of the 32 hospitals interviewed, 21 confirmed that they lead or participated in a Medicare ACO, 18 in a Medicare Shared Savings Program ACO and three in a Realizing Equity, Access, and Community Health (REACH) ACO. The remaining 11 hospitals did not participate in a Medicare ACO. Eight had no ACO experience, two participated in a commercial ACO, and one had reported experience treating ACO patients while not participating in an ACO.

Exhibit VII-1: Most Interviewees Participated in a Medicare ACO

Participate in a Medicare ACO	ACO Participation	Count
Yes	Medicare Shared Savings Program	18
	ACO REACH	3
	Total	21
No	No ACO experience	8
	Commercial ACO	2
	Experience treating ACO patients	1
	Total	11

Notes: ACO = Accountable Care Organization; REACH = Realizing Equity, Access, and Community Health.

Interviews typically included two to four interviewees who had roles such as CJR service-line coordinators, ACO representatives, population health department directors, and data analysts.

One interviewer and one notetaker conducted the interviews. Notes were taken during telephone interviews, and with the interviewee's consent, the interview was recorded to verify and enhance interview notes. ATLAS.ti software was used to code and analyze telephone interview notes.

B. How Do CJR Hospitals Interact With Medicare ACO Programs?

1. Summary of Findings

The callout box presents a summary of our findings. Results with an asterisk (*) are key findings.

- (*) Participation in a Medicare ACO influenced lower extremity joint replacement (**LEJR**) service-line care protocols and strategies to improve cost and quality of care at some hospitals.
- Hospitals made decisions to join a Medicare ACO in the context of their overall value-based care strategy.
- CJR hospitals were generally aware of a patient's ACO attribution through the electronic medical record, but most reported that their approach to care was "payer agnostic" and that all patients were treated the same regardless of ACO attribution.

2. Results

We spoke with 21 hospitals that participated in a Medicare ACO, 18 in the Medicare Shared Savings Program and three in ACO REACH. We asked these hospitals about their experience in a Medicare ACO, including decisions to participate, what it means to be part of an ACO, alignment in hospital strategies to CJR and ACO programs, and benefits and challenges to participation.

a. Decision to Join the Medicare ACO

The decision to join a Medicare ACO was often influenced by what was described as "an overall shift to value-based care" at the hospital or health system. Other reasons cited for joining a Medicare ACO include the potential to improve care coordination for patients, the desire to increase access to patient and physician education resources, and promising financial incentives. Several interviewees described having experience with value-based initiatives as a driving factor in their decision to join a Medicare ACO. Specifically, interviewees cited experience with CJR or other CMS programs, most often the Bundled Payments for Care Improvement (BPCI) Initiative or BPCI Advanced, as a reason to join or purchase a Medicare ACO. For example, one interviewee explained that CJR positively influenced the decision to continue participating in a Medicare ACO because the hospital system could leverage the technologies and funding from the CJR Model for the ACO. Another interviewee mentioned how the hospital leveraged resources from CJR to support the ACO's post-acute care (PAC) work.

> *"I would say [CJR] positively [influenced our decision on MSSP participation] because we are able to…leverage some of the technologies and…funding from the hospital to help put some of the things in place that we need for the MSSP."*
>
> ***– Manager of clinical program analytics at MSSP participant hospital***

b. Relationships With Medicare ACOs

When asked to describe what it means for their hospital to be in a Medicare ACO, respondents described increased awareness and general alignment toward value-based care. Often ACO participation was described as reflective of the hospital's commitment to value-based care and an

opportunity to prime the hospital for future success in new programs. Respondents also described ACO participation in relation to cost savings. One interviewee contextualized their hospital's ACO involvement as one of various value-based care initiatives addressing the "quadruple aim of delivering high-quality care at the lowest cost with a lens towards patient satisfaction and physician satisfaction." Another interviewee expanded on the idea, sharing that being part of a Medicare ACO means "taking accountability for the cost and quality for the attributed or assigned population while investing in the necessary technology, tools, and people for successful care coordination."

> *"[Being in a Medicare ACO means] taking accountability for the cost and quality for the attributed or assigned population. The prerequisite for success is investment in the technology, tools, and people, and also assuming […] downside risk for those activities."*
>
> ***– Director of the Population Health Service Organization at MSSP participant hospital***

We also asked interviewees about any expectations between the hospital and the Medicare ACO. Several interviewees mentioned that hospitals expect data sharing from the ACO, along with periodic discussions about potential areas of improvement and feedback on the changes already made. Several interviewees noted that the ACO expected their hospital to meet standards and protocols, including meeting contractual metrics and partaking in regularly scheduled meetings to exchange feedback. One interviewee explained that an expectation from the ACO is that the hospital will follow diagnostic standards and protocols implemented by the ACO and provide holistic care after discharge. Another interviewee shared that the hospital expected the ACO to provide information on the challenges that patients and families experience outside of the hospital, for instance, related to social determinants of health (SDOH).

c. Awareness of CJR Patient Attribution to a Medicare ACO

> *"We don't want providers to have to worry about who their patients are [insurance type]. We want patients to be treated the same, equally. Providers should be providing the best care possible for every patient they see."*
>
> ***– Manager of value-based care programs for health system at MSSP participant hospital***

Most interviewees stated that they were aware when a CJR patient was attributed to a Medicare ACO. The most common way of seeing the patient's ACO attribution was through the electronic medical record. Although they were aware of patient attribution to an ACO, interviewees noted that they were payer agnostic and that all patients in their service line receive the same care, regardless of any other program attribution. A few interviewees stated that their hospital established other communication channels through the value-based care department, care manager at the ACO, or episode coordinator from the population health department to identify patient ACO attribution. Some interviewees discussed having separate case management teams for ACO or CJR patients but did not note any differences in the care protocols used by the various case management groups.

C. How Does Medicare ACO Participation Influence the Strategies Used to Respond to CJR?

1. Summary of Findings

The callout box presents a summary of our findings. Results with an asterisk (*) are key findings.

- (*) Most hospitals agreed that CJR and Medicare ACOs have similar aims of improving quality and reducing cost.
- (*) Strategies adopted to improve the care pathway, coordinate post-acute care, and monitor patient outcomes were used by both CJR and ACOs.
- (*) Participation in a Medicare ACO or CJR resulted in an increased awareness and greater alignment toward value-based care among hospital staff.
- Participation in a Medicare ACO improved access to data used to monitor financial and quality metrics.
- Interviewees emphasized the importance of staff buy-in for achieving a value-based care culture, especially among physicians.

2. Results

a. Influence of the Medicare ACO on the LEJR Care Pathway

When reflecting on changes to the LEJR care pathway that were influenced by Medicare ACO participation, interviewees most frequently discussed leveraging ACO data such as patient medical history to guide LEJR pathway refinements and working with the ACO to improve care coordination.

ACO Data Spotlight

Interviewees stated that working with a Medicare ACO and receiving ACO data influenced care for CJR patients at their hospital, specifically allowing the hospitals improve their understanding of their patient population and interpret utilization and quality trends at their hospital to inform care redesign. ACO data about specific patients' medical history also provided CJR hospitals more background on their patients. Hospitals most frequently cited:

- **Patient medical history** including primary and specialty care – which supports the management of complex patient needs or those with comorbidities.
- **ACO patient quality data** including readmissions, hospital length of stay, and skilled nursing facility length of stay data which allows CJR providers to track patients in real time post-surgery and follow up if needed.
- **ACO population level data** including hospital utilization and quality trends and patient population characteristics as SDOH, PAC usage, and other social health needs which is used to inform care redesign at the hospital.

Almost half of interviewees stated that data sharing and access to ACO data helped them improve the monitoring of their quality and financial metrics. Several interviewees described the data received from the ACO as "timely" with some specifying that the data were received in real time

and knowing immediately when a CJR or ACO patient was readmitted. The ACO data also enhanced care coordination efforts as they allowed staff to see past medical history of CJR patients, including previous primary care and any specialty services received, which was helpful for managing comorbidities and complex patients. One interviewee reported using ACO data to "tell a story" about the patient's journey from hospital to skilled nursing facility. They used data elements such as readmissions and length of stay to highlight challenges and discuss ways to modify infrastructure.

> *"Rather than having care coordinators that were separate and in the clinics with CJR, they leveraged the Medicare Shared Savings Program care coordination program to provide care coordination for CJR patients, which helped with the alignment of the standard work."*
>
> ***– Manager of clinical program analytics at MSSP participant hospital***

About half of the interviewees discussed enhancements to care coordination at their hospital due to participation in the Medicare ACO, often involving collaboration between the hospital case management team and the ACO care management teams. Another strategy hospitals reported using was moving their care management teams under one "value-based" care management team to help coordinate care across programs. Almost half of the interviewees reported changes or refinements to discharge and PAC strategies that were influenced by participation in the ACO, including the development of preferred PAC networks and increased communication and coordination with PAC providers.

b. Alignment of Goals and Strategies Between CJR and the Medicare ACO

> *"We're pulling all the same levers. Ensuring the patient is going to the appropriate next site of care, utilizing performance networks, monitoring patient length of stay at skilled nursing facilities and monitoring the patients for readmission. The key operational levers are the same across the programs because the goals of the programs align."*
>
> ***– Assistant director of post-acute care at MSSP participant hospital***

Most interviewees felt that the goals of the two programs were similar, with both programs aiming to improve quality of care while reducing cost. Interviewees shared that common strategies were used to respond to both programs, including using data to inform care pathways; strengthening care coordination, including with PAC providers; and monitoring patient outcomes.

Interviewees described similarities in the PAC strategies used in response to the CJR Model and ACOs. A handful of interviewees described leveraging PAC strategies and resources across programs, such as using the same network of preferred PAC providers for both programs.

Interviewees said that having value-based care programs resulted in a mindset change at the hospital, for instance, increasing physicians' awareness of quality or potential economic impact. One interviewee stated, "It makes folks think about themselves as an individual that's responsible for a panel of patients. It gets physicians thinking about efficiencies that they might not have talked about in a different situation. It brings certain things like utilization to top of mind, it makes them think about social determinants of health, so I think it's helped the conversation."

> *"[The] structure forces you…to pay attention to things…that…weren't…as visible or noticeable or as accountable to, but I think we're far passed that. It's really kind of ingrained in our culture that this is the way you give good care. You know, being good stewards of resources that we have both with the patients paying what the government is paying."*
> ***– Chief Operating Officer at MSSP participant hospital***

Interviewees described hospital and system-level efforts to synchronize management of value-based care initiatives. Several hospitals reported having one singular entity, such as a population health department, that oversees all value-based care lines or using one care management entity for all value-based programs. A few interviewees mentioned standardizing their processes in response to value-based care models. One interviewee shared that an electronic management system was developed due to concurrent participation. This system allows the hospital to see patient notes and vital signs to track readmission risk. Another interviewee shared that a system-level Medicare Shared Savings Program ACO office assists the hospital orthopedic team in managing both CJR and ACO patients because they both account for total cost of care.

While most interviewees felt the goals between the Medicare ACO and CJR aligned, a handful of interviewees thought the strategies to meet the goals were unique due to the differences in the programs. Interviewees spoke of differences in patient population covered by each program and the different episode lengths, which affected the strategies used to respond.

c. Staff Buy-in and Relationships With Physicians

Most interviewees described the importance of educational efforts to promote staff buy-in for adopting value-based care. Several interviewees mentioned the importance of surgeon buy-in, specifically. In addition to hospital providers, hospitals educated other partners such as PAC providers and executive leadership. A few interviewees mentioned these efforts were led by the ACO.

> *"The level of engagement that we have had from the physician base as a result of BPCIA, CJR, and the ACO has increased dramatically."*
> ***– Manager of Bundled Payments at MSSP participant hospital***

Strategies to promote buy-in included regular meetings with providers to promote engagement, sharing data or quality metrics around value-based care with stakeholders, and streamlining value-based care programs to reduce confusion. For engaging primary care physicians and surgeons, the most common strategy adopted was meeting with stakeholders to share data and evidence-based practices. A few interviewees also discussed using the ACO to align and build relationships with primary care physicians. An interviewee from one hospital described that participation in the ACO enhanced the hospital's relationship with primary care providers because the ACO met with physicians to share best practices and data on performance metrics, quality measures, and financial reports. Another interviewee mentioned that the ACO has administrative requirements for physicians and advanced practice providers to partake in the program, including educational and advanced care training sessions and one-on-one meetings. Interviewees noted that the Medicare ACO provided stronger education initiatives and tools for physicians.

D. What Benefits and Challenges Do CJR Hospitals Participating in ACO Programs Experience?

1. Summary of Findings

- Key benefits to concurrent participation included improved access to data and similar care pathway changes.
- Identified challenges included managing two value-based care programs and understanding how the patients are attributed to each program.

2. Results

Interviewees identified several benefits to concurrent participation in the CJR Model and a Medicare ACO. Several mentioned that the CJR program is small compared with their Medicare ACO program, which has a much larger attributed population, and concurrent participation allows the CJR program to benefit from the investments and attention to value-based care that come with ACO participation.

Interviewees also identified increased data access as a key benefit, describing the ability to access and analyze more data because of concurrent participation in the CJR Model and the ACO. Hospitals leveraged these data to look at trends in patient health conditions and identify opportunities for quality improvement.

Interviewees identified a few challenges to concurrent participation in both models, including issues with financial management and administrative burdens. They discussed the challenges of managing the overlap of several value-based care programs at the hospital, noting struggles with allocating savings between the two programs at the hospital when they share resources such as care coordinators and difficulties determining the net financial impact without knowing how the multiple programs interact with each other. Other interviewees indicated that it was difficult to proactively manage CJR patients because it was unclear which episodes would be attributed to the ACO and which would count as a CJR episode as attribution rules varied ACO type and participation year.

Interviewees at several hospitals mentioned challenges working with ACO and CJR data. Some hospitals reported difficulty in having staff with experience and knowledge to interpret the data in a useful way, and some relied on outside consultants. Challenges with the data itself include delayed performance and outcomes data from CMS and the difficulty managing data from multiple programs. One hospital shared that they receive less data from CJR than the ACO, which makes it harder to be more proactive for CJR.

E. Do CJR Hospitals That Do Not Participate in a Medicare ACO Have Any Experience with ACOs?

1. Summary of Findings

- The majority of hospital representatives interviewed described little to no interaction with ACOs, and most described that CJR providers were not aware of beneficiary attribution to an ACO.
- When asked about the reasons for not joining an ACO, interviewees described local market changes at the hospital, financial concerns, risks of the program outweighing the benefits, and a lag in getting real-time data.

2. Results

We spoke to representatives at 11 CJR hospitals that do not participate in a Medicare ACO; however, two of these hospitals reported participating in a non-Medicare commercial ACO. Other hospitals in this group may care for patients attributed to Medicare ACO or have no interaction with any ACO. Due to the varying options for ACO participation or affiliation, hospital interviewees were asked to describe the overlap between the CJR Model and ACO and whether the ACO influenced the CJR Model at their hospital. We also asked interviewees about a variety of topics related to the ACO, such as influence on LEJR care pathways, conversations related to joining the ACO, and hospital or service-line awareness of beneficiary attribution to an ACO.

> *"[Joining a commercial ACO] was a good way to diversify, not just traditional sources of revenues coming in, but it gave us an additional one [source of revenue] that we had a really good chance at, so we stayed."*
>
> ***– ACO Executive Director at hospital that previously participated in a Medicare NextGen ACO***

> *"[The] pathway has always been to do what's best for the patient and get them home or to the next level of care, without particular guidance to do this or don't do this."*
>
> ***– Director of Rehabilitation Services at hospital that previously participated in a Medicare NextGen ACO***

Most hospitals interviewed had little to no interaction with ACO leadership, structures, or processes and described that CJR providers were not aware of beneficiary attribution to an ACO. When asked about the reasons for not joining an ACO, interviewees described local market changes at the hospital, financial concerns, risks of the program outweighing the benefits, and a lag in getting performance and target pricing data in a timely manner.

Only three hospitals participated in a commercial ACO or reported caring for patients attributed to a Medicare ACO. These hospital representatives provided varying perspectives on the program overlap and the ACO's influence on the CJR Model. However, when asked about the care that patients received, it emerged that patients attributed to an ACO may receive care beyond what is solely provided from the CJR Model. One interviewee explained that the extra efforts from an ACO are considered "above and beyond this optimized process [efficient CJR pathways at the hospital]." Another interviewee explained that the hospital used and benefited

from the preferred PAC network that was created by the commercial ACO as part of its response to the CJR Model, which allowed the hospital staff to have leverage and input on patient care plans.

> *"There is a long list of individuals that are included in the decision to join an ACO, including the clinical institute the hospital system is a part of, hospital executives, finance executives, program managers, and direct care providers."*
>
> ***– Executive Director for Orthopedic, Sports Medicine, and Spine at non-ACO-affiliated hospital***

VIII. Considerations

We examined the impact of the Comprehensive Care for Joint Replacement (CJR) Model on episode-of-care payments, utilization of post-acute care services, and quality of care for lower extremity joint replacements (LEJRs) using varied data sources and analytic approaches. This chapter highlights factors that may have influenced the observed results and readers should consider when interpreting the findings.

Acronyms	
ASC	ambulatory surgical centers
BPCI	Bundled Payments for Care Improvement
CJR	Comprehensive Care for Joint Replacement
CMS	Centers for Medicare & Medicaid Services
DiD	difference-in-differences
LEJR	lower extremity joint replacement
MA	Medicare Advantage
MSA	metropolitan statistical area
PY	performance year
THA	total hip arthroplasty
TKA	total knee arthroplasty

Addition of outpatient episodes. Starting in Performance Year 6 (PY 6), the Centers for Medicare & Medicaid Services (CMS) included outpatient total knee arthroplasty (TKA) and total hip arthroplasty (THA) as LEJR episodes under the CJR Model. Controlling for the surgery setting presented a challenge for our difference-in-differences (DiD) analyses and doing so could have led to biased impact estimates. In statistical terms, it would have violated the exogeneity assumption that underlies ordinary least squares. Because of this concern, we did not risk-adjust for the surgical setting. As a result, our DiD impact estimates captured the combined effects of any influence on outcomes of the model that may affect the setting of the LEJR along with any changes in outcomes the model caused directly through care transformation within the inpatient or outpatient setting. Separating the two possible types of effects was not feasible, even though it would have been useful and informative, so readers should consider our DiD impact estimates as the overall effects of the CJR Model, including effects related to changes in the surgery setting.

Changes in included hospitals. As part of the CJR Model extension, CMS updated the list of hospitals exempted from the model due to rural or low-volume status to include hospitals certified by CMS as rural as of October 4, 2020.[90] Many rural CJR hospitals for whom participation had been mandatory since the start of the model no longer participate in the model in PY 6. Two potential effects could have resulted from this change. First, because hospitals had to apply for rural status, this change in hospital participation could have potentially biased estimates of model impact due to non-random changes to CJR hospital composition if unobserved characteristics are correlated with the decision to apply. Second, removing rural CJR hospital participants may also have affected generalizability of the findings because the CJR Model effects may have differed for rural areas. Therefore, we cannot easily extrapolate the impacts observed in PY 6 to certain areas of the country.

Growth in Medicare Advantage (MA) enrollment. MA enrollment grew steadily during the past 20 years. During the period of the CJR Model, nationally 33% of Medicare beneficiaries enrolled in MA plans in 2016 and grew to 48% for Calendar Year 2022, which coincided with the last 4

[90] Previously, CMS determined rural and low-volume hospital status as of February 1, 2018. Between February 1, 2018, and October 4, 2020, additional hospitals applied for and CMS granted rural or low-volume status. This designation resulted in their exclusion from CJR starting in PY 6. See the CJR Three-Year Extension Final Rule, Section D.2, for more details: https://www.federalregister.gov/documents/2021/05/03/2021-09097/medicare-program-comprehensive-care-for-joint-replacement-model-three-year-extension-and-changes-to.

quarters of PY 6.[91] With the growth in MA that occurred throughout the CJR Model, CJR and control metropolitan statistical areas (MSAs) have varying levels of MA penetration that may have biased our estimates. However, we did not see any notable changes in patient characteristics between the CJR and control groups in our analysis described in **Chapter VI: Analysis of Potential Unintended Consequences of the CJR Model**.

Participation in other value-based payment models. CJR and control hospitals participate or have participated in other value-based payment models, which could affect their CJR Model performance. These models include the Bundled Payments for Care Improvement (BPCI) initiative and its successor, BPCI Advanced. As discussed in the CJR Model evaluation's Fifth Annual Report,[92] the control group has a high level of hospital participation in the BPCI Advanced model for LEJR clinical episodes. Although we controlled for BPCI Advanced alignment in the control group, the fact that CJR participant hospitals cannot participate in BPCI Advanced for LEJR related episodes could have resulted in an underestimate of the overall effect of the CJR Model.

In the CJR Model evaluation's Fifth Annual Report, we noted that 50% of control hospitals belonged to the same health system as at least one mandatory CJR hospital. This continued to be an important factor when considering model impacts in PY 6. If CJR hospitals in these health systems shared care practices with control hospitals, the control group could have become contaminated, likely resulting in an underestimate of the CJR Model.

Addition of Medicare coverage for ambulatory surgical centers (ASCs). Medicare began to allow the performance of knee and hip LEJRs at ASCs in 2020 and 2021, respectively. However, the CJR Model does not include LEJRs performed at ASCs so we did not include ASC LEJRs in any of our analyses. The share of ASC LEJRs has grown slowly but steadily and comprised 6–8% of TKAs and 3–5% of THAs across CJR and control MSAs in PY 6. Because Medicare did not cover ASC LEJRs before 2020, which includes our baseline period, our DiD design does not naturally account for this policy change. If the surgeons in CJR MSAs performed more or fewer LEJRs in the ASC in response to CJR incentives, our DiD impact estimates did not capture this behavior. Moving forward, we will continue to monitor the share of LEJRs performed in the ASC setting and modify our analytic approach if necessary.

[91] We obtained Medicare Advantage statistics from a Kaiser Family Foundation analysis of CMS Medicare Advantage Enrollment Files: Ochieng, N., Fuglesten Biniek, J., Freed, M., Damico, A., & Neuman, T. (2023). *Medicare Advantage in 2023: enrollment update and key trends*. https://www.kff.org/medicare/issue-brief/medicare-advantage-in-2023-enrollment-update-and-key-trends/

[92] Centers for Medicare & Medicaid Services. (2023). *Comprehensive Care for Joint Replacement Model - Fifth Annual Report.* https://www.cms.gov/priorities/innovation/data-and-reports/2023/cjr-py5-annual-report

Appendix A: List of Acronyms and Glossary of Terms

Exhibit A-1: List of Acronyms

Acronym	Meaning
ACH	Acute Care Hospital
ACO	Accountable Care Organization
ADLs	Activities of Daily Living
AHA	American Hospital Association
APM	Alternative Payment Model
ASC	Ambulatory Surgical Center
BPCI	Bundled Payments for Care Improvement
BPCI-A	Bundled Payments for Care Improvement Advanced
CHF	Congestive Heart Failure
CI	Confidence Interval
CJR	Comprehensive Care for Joint Replacement
CMMI	Center for Medicare & Medicaid Innovation
CMS	Centers for Medicare & Medicaid Services
COVID-19	Coronavirus Disease 2019
CPT	Current Procedural Terminology
CY	Calendar Year
DDD	Difference-in-Difference-in-Differences
DiD	Difference-in-Differences
DME	Durable Medical Equipment
DSH	Disproportionate Share Hospital
ED	Emergency Department
EMR	Electronic Medical Record
ESRD	End-Stage Renal Disease
FFS	Fee-for-Service
FY	Fiscal Year
HCC	Hierarchical Condition Category
HH	Home Health
HHA	Home Health Agency
IP	Inpatient
IPO	Inpatient Only
IPPS	Inpatient Prospective Payment System
IRF	Inpatient Rehabilitation Facility
LCI	Lower confidence interval
LEJR	Lower Extremity Joint Replacement
LOS	Length of Stay

Acronym	Meaning
LTCH	Long-Term Care Hospital
MA	Medicare Advantage
MCC	Major Complication or Comorbidity
MDS	Minimum Data Set
MSA	Metropolitan Statistical Area
MS-DRG	Medicare Severity-Diagnosis Related Group
MSSP	Medicare Shared Savings Program
NPRA	Net Payment Reconciliation Amount
OLS	Ordinary least squares
OP	Outpatient
OT	Occupational Therapy
PAC	Post-Acute Care
PCP	Primary Care Provider
PDGM	Patient Driven Groupings Model
PDP	Post Discharge Period
PDPM	Patient Driven Payment Model
PEP	Post Episode Period
PGP	Physician Group Practice
PHE	Public Health Emergency
PRO	Patient-Reported Outcomes
PSW	Propensity Score Weighting
PT	Physical Therapy
PY	Performance Year
RIF	Research Identifiable File
SDOH	Social Determinants of Health
SNF	Skilled Nursing Facility
SNH	Safety Net Hospital
THA	Total Hip Arthroplasty
TKA	Total Knee Arthroplasty
UCI	Upper confidence interval
VBC	Value-Based Care
VBP	Value-Based Payments

Exhibit A-2: Glossary of Terms

Term	Definition
Acute care hospital (ACH)	A health care facility that provides inpatient medical care and other related services for acute medical conditions or injuries.
Ambulatory surgical center (ASC)	A health care facility that provides surgical care to patients not requiring hospitalization or services exceeding 24 hours.
Anchor hospitalization or procedure	The inpatient hospitalization or outpatient procedure that triggers the start of the episode of care.
Baseline time period	The period of time that precedes the intervention period as a basis for comparison in the difference-in-differences statistical technique. The baseline period includes episodes that were initiated from 2012 to 2014 and that ended between April 1, 2012, and March 31, 2015.
Beneficiary incentive	Programmatic flexibility available to hospitals participating in the CJR Model. This allows participating hospitals to offer patients certain incentives not tied to the standard provision of health care, as long as it supports a clinical goal.
Bundle	The services provided during the episode that are linked for payment purposes.
CJR collaborator	Medicare-enrolled providers and suppliers engaged in caring for CJR beneficiaries that enter into sharing agreements with a participant hospital. Collaborators may be a SNF, HHA, LTCH, IRF, physician, non-physician practitioner, provider or supplier of outpatient therapy services, PGP, non-physician provider group practice, ACO, hospital, or critical access hospital.
CJR sharing arrangement	A financial arrangement between a participant hospital and a CJR collaborator for the sole purpose of making gainsharing payments or alignment payments under the CJR Model.
Effective discount percentage	The effective discount percentage serves as Medicare's portion of the savings. A 3% effective discount percentage is used to set the prospective quality-adjusted target price. The effective discount percentage used at reconciliation varies based on the hospital's quality performance in the year and whether the hospital's average episode payment falls above or below its quality-adjusted target price. For hospitals receiving reconciliation payments, the effective discount percentages are: 1.5% for "excellent" quality, 2% for "good" quality, and 3% for "acceptable" quality. Hospitals with "below acceptable" quality are ineligible to receive reconciliation payments. For hospitals with repayment responsibility in PY 6, a 1.5 percentage point reduction to the applicable discount percentage was made for participant hospitals with "good" quality performance and a 3-percentage point reduction to the applicable discount factor was made for participant hospitals with "excellent" quality performance.
Episode benchmark price	The episode benchmark price represents the expected episode payments if treatment patterns and patient mix do not change from previous historical spending for LEJR episodes. For Performance Year 6, the episode benchmark price is based solely on regional amounts and include both outpatient TKA and THA. Additionally, benchmark prices are adjusted for age indicators, hierarchical category condition counts, and dual eligibility for Medicaid beginning in Performance Year 6. The product of the episode benchmark price and the effective discount percentage equals the quality-adjusted target price.

Term	Definition
Episode of care	A CJR episode of care is triggered by the admission of an eligible Medicare fee-for-service beneficiary to a hospital paid under IPPS for an inpatient hospitalization or an outpatient LEJR procedure that results in a discharge paid under MS-DRG 469 (major joint replacement or reattachment of lower extremity with MCC), 470 (major joint replacement or reattachment of lower extremity without MCC), MS-DRG 521 (hip replacement with principal diagnosis of hip fracture with MCC), or MS-DRG 522 (hip replacement with principal diagnosis of hip fracture without MCC), and ends 90 days after discharge from the anchor hospitalization or the outpatient procedure.
Gainsharing payment	A payment from a participant hospital to a CJR collaborator made pursuant to a CJR sharing arrangement. A gainsharing payment may be composed of reconciliation payments, internal cost savings, or both.
Inpatient-only (IPO) list	A list of procedures that are covered by Medicare only when provided in the inpatient setting.
Internal cost savings (ICS)	The measurable, actual, and verifiable cost savings realized by the CJR-participating hospital resulting from care redesign undertaken by the hospital in connection with providing items and services to CJR Model beneficiaries. Internal cost savings does not include savings realized by any individual or entity that is not a CJR participant hospital.
Metropolitan statistical area (MSA)	A geographical region that is associated with a core urban area and has a population of at least 50,000.
Net Payment Reconciliation Amount (NPRA)	The aggregate quality-adjusted target price minus the total dollar amount of Medicare fee-for-service payments for items and services included in the bundle, adjusted by stop gain or stop loss limits, if applicable.
Outpatient (OP) department	A hospital-based care setting for procedures covered by Medicare through the Outpatient Prospective Payment System. The 2-midnight rule provides guidance regarding the classification of inpatient or outpatient procedures.
Post-acute care (PAC)	Rehabilitation and palliative care services received by the beneficiary from IRFs, SNFs, HHAs, or LTCHs following a hospitalization.
Post-discharge home visit waiver	A waiver available to hospitals participating in the CJR Model which waives the direct supervision requirement for home visits. CJR beneficiaries may receive up to nine home visits per episode by licensed clinical staff paid under the Medicare Physician Fee Schedule.
Post-discharge period (PDP)	The 90-day period after discharge, which starts on the day of the anchor hospitalization discharge.
Post-episode care	Care that occurs after the 90-day post-discharge period under the CJR Model.
Quality-adjusted target price	The quality-adjusted target price for episodes in PY 6 is based solely on the regional historical episode payment for LEJR episodes beginning in Calendar Year 2019. At the beginning of the performance year, it is assumed that the hospital's composite quality score falls in the "acceptable" range, which is then modified to reflect the hospital's actual composite quality score at reconciliation. There are separate quality-adjusted target prices to account for MS-DRG and hip fracture status. Additionally, target prices are adjusted for age indicators, hierarchical category condition counts, and dual eligibility for Medicaid.
Reconciliation payment	A retrospective payment between Medicare and a CJR participant hospital at the end of a performance year. If total fee-for-service payments for its episodes during a performance year are less than the aggregate quality-adjusted target price, Medicare makes a payment to a CJR participant hospital. If total fee-for-service payments for a CJR participant hospital's episodes are more than its aggregate quality-adjusted target price, the hospital repays the difference to Medicare.

Term	Definition
Related items and services	Episode-related items and services paid under Medicare Part A or Part B that are included in the bundle after exclusions are applied. These episode-related items and services include physicians' services, inpatient hospital services (including readmissions with certain exceptions discussed in the Final Rule), inpatient psychiatric facility services, LTCH services, IRF services, SNF services, HHA services, hospital outpatient services, outpatient therapy services, clinical laboratory services, DME, Part B drugs, and hospice.
Risk adjustment	A statistical process to adjust claims-based outcomes and ADL measures to consider differences at the patient, episode, hospital, state, and MSA level that are related to the measures of interest. Without adequate risk adjustment, providers treating a sicker or more service-intensive patient mix would have worse outcomes than otherwise comparable providers serving healthier patients.
Stop-loss/Stop-gain limits	Adjustments included in the NPRA calculation vary by performance year. The stop-loss limit is the maximum amount a hospital will have to repay to CMS, and the stop-gain limit is the maximum amount that a hospital will receive from CMS as a reconciliation payment. They are based on a percentage of the quality-adjusted target price. The stop-loss and stop-gain limits are 20% in PY 6.
Telehealth waiver	A waiver available to hospitals participating in the CJR Model. Under this waiver, CMS allows Medicare coverage of telehealth services furnished to eligible beneficiaries regardless of their geographic region. Further, the originating site requirement is waived for eligible beneficiaries receiving telehealth services from their homes or places of residence.
Three-day hospital stay waiver	A waiver available to hospitals participating in the CJR Model. Under this waiver, CMS waives the three-day hospital stay requirement for Part A skilled nursing facility coverage.

Appendix B: Data and Methods

In this chapter, we summarize the data and methods used to evaluate the Comprehensive Care for Joint Replacement (CJR) Model and generate the results presented in this Sixth Annual Report.

A. Secondary Data Sources

Exhibit B-1: Secondary Data Sources

Data source	Date range	Dataset contents	Use
American Hospital Association (AHA) survey data	2021	AHA Annual Survey Database for FY 2021 is a hospital database for peer comparisons, market analysis, and health services research.	Used to Identify ACO participating Hospitals for telephone interviews.
Area Health Resource Files (AHRF)	2015-2016 (Data are from 2012-2014)	County-level data aggregated to the MSA level. Variables include Medicare Advantage penetration, average Medicare beneficiary HCC score, dual eligible percentage, population per square mile, geography, and supply of health care facilities (SNF beds, long-term care hospital beds) and health care professionals (primary care physicians, orthopedic surgeons, nurse practitioners or physician assistants, specialists).	Used to control for MSA Medicare Advantage penetration in the patient survey analysis.
Bundled Payments for Care Improvement Advanced (BPCI-A) programmatic participant data	Intervention	Identifies health care providers (hospitals, physicians, and physician practice groups) that are participating in the BPCI-A model, the time period of participation, and the episodes for which they are participating.	Used to identify LEJR discharges in the control group that are assigned to BPCI-A participants for risk adjustment.
Bundled Payments for Care Improvement (BPCI) programmatic participant data	Baseline and intervention	Identifies health care providers (hospitals, PAC providers, physicians, and physician practice groups) that are participating in the BPCI initiative, the time period of participation, and the models and episodes for which they are participating.	Used to identify LEJR discharges that are assigned to BPCI participants for exclusion. Used to identify hospitals as past BPCI LEJR participants for risk adjustment.
Comprehensive Care for Joint Replacement (CJR) programmatic data	Intervention	List of CJR participant hospitals, as well as their PY 1, PY 2, PY 3 PY 4, PY 5, and PY 6 quality-adjusted target prices, reconciliation (NPRA), and hospital quality data.	Used to identify CJR participating hospitals, hospitals that continued mandatory participation in PY 3, their start and end dates in the CJR Model, their quality performance, and their reconciliation payments or repayment responsibility. Used total reconciliation payments and repayments to CMS to calculate savings to Medicare and investigate the distribution of NPRA.

Data source	Date range	Dataset contents	Use
FY Acute IPPS Final Rule data files	FY 2016, FY 2018, FY 2021	On an annual basis, CMS sets acute care hospital IPPS payment rates. Data files include FY hospital-level information on provider identification number, bed count, medical residents per 1,000 beds, average daily census, disproportionate share hospital patient percentage, uncompensated care payment per claim, Medicare days as a percent of total inpatient days, and section 401 (rural) status.	Used to risk adjust for acute care IPPS hospital characteristics. Used to identify section 401 hospitals (rural designation) located in control group hospitals to exclude from the PY 6 sample.
CMS Master Data Management (MDM)	Baseline and Intervention	Provider- and beneficiary- level information on participation in CMS Innovation Center payment demonstration programs. Includes beneficiary ID, program ID, and start and end dates of participation.	Used to identify beneficiaries involved in Pioneer, Next Generation, and MSSP ACO programs and control for their participation in our analyses. Used to apply the ACO exclusion for episodes starting on or after July 1, 2017 (MSSP track 3, Comprehensive End-Stage Renal Disease Care Model with downside risk, and Next Generation). For PY 6 through PY 8, SSP Enhanced Track episodes were excluded.
Medicare FFS beneficiary enrollment data	Baseline and Intervention	Enrollment data (from Common Medicare Enrollment and Medicare Beneficiary Summary File) provide beneficiary Medicare Parts A and B eligibility information.	Enrollment data were used to confirm beneficiary eligibility and provide beneficiary characteristics for analyses (for example, risk adjustment models). Enrollment data were used to measure the change in case-mix of CJR and control group patients between the baseline and the intervention periods.
Medicare FFS claims	Baseline and Intervention	Medicare Parts A and B claims data (from monthly Medicare claims [TAP] files) provide claims for different services received during the anchor hospitalization and post-discharge period (for example, dates and types of service). A minimum of 3-month claims run out was used for episodes included in this report.	Claims were used to: 1) create the CJR episodes, describe service use, and create risk adjustment (for example, beneficiary prior utilization, HCC score, COVID-19 diagnosis) and outcome variables (for example, unplanned readmissions, emergency department visits, and number of days or visits in each PAC setting); 2) identify TKA and THA procedures in the hospital outpatient departments and ambulatory surgical centers in CJR and control markets for descriptive analyses and create outpatient TKA and THA episodes; and 3) sample participants for primary data collection (patient survey, telephone interviews).

Data source	Date range	Dataset contents	Use
MSSP ACO Provider Research Identifiable File (RIF)	Baseline, 2021, and 2022	The Shared Savings Provider RIF file years 2013, 2014, 2021, and 2022 is used to identify hospitals at the CCN level participating in the SSP.	Used to compile Medicare ACO participation data among CJR and control hospitals.
MDS 3.0 data	Baseline and Intervention	The MDS is a comprehensive assessment instrument administered by nursing staff to all Medicare beneficiaries when they are admitted to a Medicare-certified SNF, at discharge, as well as on days 5, 14, 30, 60, 90, and quarterly, thereafter. The MDS collects information on patients' demographics, history and diagnoses, skin conditions, medications, care management, restraint use, preferences for routine and activities, and functional, sensory, cognitive, neuro or emotional, bladder, bowel, swallowing or nutritional, and pain status. A minimum 6 month run out of MDS data was used for episodes included in this report.	MDS data were used to identify patients who were in a SNF or long-term nursing facility during the 6 months preceding the episode which was used for risk-adjustment and to measure the change in case-mix of CJR patients and patients in the control group between the baseline and the intervention periods.
Medicare standardized payments	Baseline and Intervention	Medicare standardized payments for 100% of Parts A and B claims received via the Integrated Data Repository. Produced by a CMS contractor.	Used to create Medicare standardized paid amounts (Parts A and B) and allowed standardized payment amounts, including beneficiary out-of-pocket amounts. Used to estimate the impacts of the CJR Model on total episode and service-level payments.
Next Generation ACO Provider Research Identifiable File (RIF)	2021	Information on Next Generation ACO NGACO provider-level participation and related data. The Next Generation ACO Provider RIF files year 2021 is used to identify hospitals at the CCN level participating in the Next Generation Program.	Used to compile Medicare ACO participation data among CJR and control hospitals.
Pioneer ACO Model Provider Research Identifiable File (RIF)	Baseline	Information on Pioneer ACO provider-level participation and related data. The Pioneer ACO Provider RIF files years 2012, 2013, and 2014 is used to identify hospitals at the CCN level participating in the Pioneer program.	Used to compile Medicare ACO participation data among CJR and control hospitals.
Provider of Services (POS) file	December 2016	Information on Medicare-approved facilities, including provider identification number, ownership status, size, medical school affiliation, and staffing.	Used to identify and characterize acute care hospitals actively engaged in Medicare for risk adjustment and descriptive analyses.

Data source	Date range	Dataset contents	Use
USAFacts	2020 and later	Information on COVID-19 infection rates at the national, state and county-levels. Aggregated from individual federal, state, and local data systems.	Used to create risk adjustment variables for *select* analyses that account for county-level COVID-19 infection rates at the time of the episode start date. Infection rates were linked to the CJR and control hospital's county of residence. At the end of the Public Health Emergency during PY 6, some counties stopped reporting reliable data. Thus, we exclude these risk adjustment variables from most of our analyses. In our analyses we did include them, we used a version that was capped at the max rate to account for potential lasting shocks to the local health care system.

Notes: ACO = Accountable Care Organization, AHA = American Hospital Association, AHRF = Area Health Resource Files, BPCI = Bundled Payments for Care Improvement, BPCI-A = Bundled Payments for Care Improvement Advanced, CCN = CMS certification number, CJR = Comprehensive Care for Joint Replacement, CMS = Centers for Medicare & Medicaid Services, COVID-19 = coronavirus disease 2019, FFS = fee-for-service, FY = fiscal year, HCC = hierarchical condition category, ID = identification, IPPS = Inpatient Prospective Payment System, LEJR = lower extremity joint replacement, MDM = Master Data Management, MDS = Medicare Minimum Data Set 3.0, MSA = metropolitan statistical area, MSSP = Medicare Shared Savings Programs, NPRA = net payment reconciliation amount, PAC = post-acute care, POS = provider of services, PPS = prospective payment system, PSW = propensity score weight, PY = performance year, RIF = Research Identifiable File, SNF = skilled nursing facility, SSP = shared savings program, THA = total hip arthroplasty, TKA = total knee arthroplasty.

B. Study Population

This section defines the CJR and control group populations, explains the weights used in the mandatory analyses to account for differences in sampling probabilities, and outlines the additional eligibility criteria for hospitals and episodes.

1. Defining the CJR and Control Group Populations

At the start of the CJR Model in 2016, CMS selected MSAs eligible for CJR participation based on a stratified sampling methodology. CMS stratified MSAs into eight strata based on historical wage-adjusted episode payments and population size. Within each stratum, MSAs were randomly selected to participate in the CJR Model (n=67 MSAs). This design allowed for a control group of hospitals in MSAs that were eligible but not selected by CMS to participate in the CJR Model (n=104 MSAs). These MSAs represented what would have happened in CJR-type markets if the model was never implemented (that is, the *counterfactual*). To be included in the CJR Model and in our analysis, hospitals had to be acute care hospitals paid under the IPPS that performed LEJRs for Medicare beneficiaries.

In 2018, the 67 original CJR MSAs were ranked by average historical wage-adjusted episode payment, and CMS required the top 34 MSAs with the highest payments to continue participation in the model (*mandatory* MSAs) while giving hospitals in the bottom 33 MSAs a one-time opportunity to opt in (*voluntary* MSAs). This change reduced mandatory participation by about half by allowing all CJR hospitals in the 33 low-payment MSAs and rural or low-volume CJR hospitals in the 34 high-payment MSAs a one-time opportunity to remain in the model. As discussed in **Chapter I: CJR Model Background**, starting in PY 6, CMS required all CJR participating hospitals in voluntary MSAs to halt their participation in the model. Also starting in PY 6, low-volume or rural CJR hospitals participating in mandatory MSAs could no longer participate in the model.

This report covers PY 6 of the CJR Model, from October 1, 2021, to December 31, 2022. Our analyses focused on episodes from hospitals that were mandated to participate in PY 6. As such, we excluded rural and low-volume hospitals in the mandatory MSAs that previously could opt in to continue participation in CJR because they were no longer allowed to remain in the model in PY 6. Low-volume hospitals were defined as hospitals that had less than 20 episodes over a 3-year historical period (2012 to 2014), and we identified rural hospitals using the FY 2019 IPPS data (Section 401 hospitals). Starting in PY 6, a considerable number of CJR hospitals for which participation had been mandatory since the beginning of the model were no longer eligible to be in the model. Of the 395 mandatory CJR hospitals in PYs 3–5, 72 no longer participated in the model in PY 6 due to an updated rural designation status.

Exhibit B-2 provides the names of the CJR and control group MSAs included in the original CJR Model and in our analysis of PY 6.

Exhibit B-2: CJR and Control Group MSAs Included in the Original Model and PY 6

CJR		Control	
CBSA ID	MSA name, state	CBSA ID	MSA name, state
10420	Akron, OH*	10180	Abilene, TX
10740	Albuquerque, NM	10580	Albany-Schenectady-Troy, NY
11700	Asheville, NC*	10900	Allentown-Bethlehem-Easton, PA-NJ*
12020	Athens-Clarke County, GA	11100	Amarillo, TX*
12420	Austin-Round Rock, TX*	11260	Anchorage, AK
13140	Beaumont-Port Arthur, TX*	12060	Atlanta-Sandy Springs-Roswell, GA
13900	Bismarck, ND	12700	Barnstable Town, MA*
14500	Boulder, CO	13460	Bend-Redmond, OR
15380	Buffalo-Cheektowaga-Niagara Falls, NY	13820	Birmingham-Hoover, AL*
16020	Cape Girardeau, MO-IL	14260	Boise City, ID
16180	Carson City, NV	14460	Boston-Cambridge-Newton, MA-NH
16740	Charlotte-Concord-Gastonia, NC-SC	14540	Bowling Green, KY*
17140	Cincinnati, OH-KY-IN*	15940	Canton-Massillon, OH
17860	Columbia, MO	15980	Cape Coral-Fort Myers, FL*
18580	Corpus Christi, TX*	16060	Carbondale-Marion, IL*
19500	Decatur, IL	16300	Cedar Rapids, IA
19740	Denver-Aurora-Lakewood, CO	16620	Charleston, WV
20020	Dothan, AL	16700	Charleston-North Charleston, SC
20500	Durham-Chapel Hill, NC	16860	Chattanooga, TN-GA*
22420	Flint, MI	16980	Chicago-Naperville-Elgin, IL-IN-WI*
22500	Florence, SC	17020	Chico, CA
23540	Gainesville, FL*	17780	College Station-Bryan, TX
23580	Gainesville, GA	17900	Columbia, SC*
24780	Greenville, NC*	17980	Columbus, GA-AL
25420	Harrisburg-Carlisle, PA*	18140	Columbus, OH
26300	Hot Springs, AR*	19100	Dallas-Fort Worth-Arlington, TX*
26900	Indianapolis-Carmel-Anderson, IN	19380	Dayton, OH*
28140	Kansas City, MO-KS	19660	Deltona-Daytona Beach-Ormond Beach, FL*
28660	Killeen-Temple, TX*	19820	Detroit-Warren-Dearborn, MI*
30700	Lincoln, NE	20260	Duluth, MN-WI
31080	Los Angeles-Long Beach-Anaheim, CA*	20740	Eau Claire, WI
31180	Lubbock, TX*	22020	Fargo, ND-MN
31540	Madison, WI	22520	Florence-Muscle Shoals, AL*
32820	Memphis, TN-MS-AR*	22900	Fort Smith, AR-OK
33100	Miami-Fort Lauderdale-West Palm Beach, FL*	23060	Fort Wayne, IN
33340	Milwaukee-Waukesha-West Allis, WI	23420	Fresno, CA
33700	Modesto, CA	24340	Grand Rapids-Wyoming, MI
33740	Monroe, LA*	24580	Green Bay, WI
33860	Montgomery, AL*	24860	Greenville-Anderson-Mauldin, SC*

CJR		Control	
CBSA ID	**MSA name, state**	**CBSA ID**	**MSA name, state**
34940	Naples-Immokalee-Marco Island, FL	25060	Gulfport-Biloxi-Pascagoula, MS*
34980	Nashville-Davidson--Murfreesboro--Franklin, TN	25540	Hartford-West Hartford-East Hartford, CT
35300	New Haven-Milford, CT*	25620	Hattiesburg, MS*
35380	New Orleans-Metairie, LA*	25940	Hilton Head Island-Bluffton-Beaufort, SC*
35620	New York-Newark-Jersey City, NY-NJ-PA*	26140	Homosassa Springs, FL*
35980	Norwich-New London, CT	26420	Houston-The Woodlands-Sugar Land, TX*
36260	Ogden-Clearfield, UT	26580	Huntington-Ashland, WV-KY-OH
36420	Oklahoma City, OK*	26620	Huntsville, AL*
36740	Orlando-Kissimmee-Sanford, FL*	26980	Iowa City, IA
37860	Pensacola-Ferry Pass-Brent, FL	27140	Jackson, MS*
38300	Pittsburgh, PA*	27860	Jonesboro, AR*
38940	Port St. Lucie, FL*	27900	Joplin, MO
38900	Portland-Vancouver-Hillsboro, OR-WA	29180	Lafayette, LA*
39340	Provo-Orem, UT*	29200	Lafayette-West Lafayette, IN
39740	Reading, PA*	29340	Lake Charles, LA*
40980	Saginaw, MI	29420	Lake Havasu City-Kingman, AZ
41860	San Francisco-Oakland-Hayward, CA	29460	Lakeland-Winter Haven, FL*
42660	Seattle-Tacoma-Bellevue, WA	29620	Lansing-East Lansing, MI
42680	Sebastian-Vero Beach, FL	30460	Lexington-Fayette, KY*
43780	South Bend-Mishawaka, IN-MI	30620	Lima, OH*
41180	St. Louis, MO-IL	30780	Little Rock-North Little Rock-Conway, AR
44420	Staunton-Waynesboro, VA	31140	Louisville/Jefferson County, KY-IN*
45300	Tampa-St. Petersburg-Clearwater, FL*	31420	Macon, GA*
45780	Toledo, OH*	31700	Manchester-Nashua, NH
45820	Topeka, KS	33460	Minneapolis-St. Paul-Bloomington, MN-WI
46220	Tuscaloosa, AL	34820	Myrtle Beach-Conway-North Myrtle Beach, SC-NC
46340	Tyler, TX*	34900	Napa, CA
48620	Wichita, KS	35840	North Port-Sarasota-Bradenton, FL*
		36100	Ocala, FL
		36540	Omaha-Council Bluffs, NE-IA
		37900	Peoria, IL
		37980	Philadelphia-Camden-Wilmington, PA-NJ-DE-MD*
		38060	Phoenix-Mesa-Scottsdale, AZ
		38860	Portland-South Portland, ME
		39300	Providence-Warwick, RI-MA
		39460	Punta Gorda, FL*
		39580	Raleigh, NC
		40140	Riverside-San Bernardino-Ontario, CA*
		40220	Roanoke, VA

CJR		Control	
CBSA ID	MSA name, state	CBSA ID	MSA name, state
		40340	Rochester, MN
		40380	Rochester, NY
		40900	Sacramento--Roseville--Arden-Arcade, CA
		41500	Salinas, CA
		41620	Salt Lake City, UT*
		41740	San Diego-Carlsbad, CA
		41940	San Jose-Sunnyvale-Santa Clara, CA
		41980	San Juan-Carolina-Caguas, PR
		42200	Santa Maria-Santa Barbara, CA
		42220	Santa Rosa, CA
		42340	Savannah, GA
		43340	Shreveport-Bossier City, LA*
		43620	Sioux Falls, SD
		44060	Spokane-Spokane Valley, WA
		44100	Springfield, IL
		44180	Springfield, MO
		41100	St. George, UT
		46060	Tucson, AZ
		46140	Tulsa, OK
		46520	Urban Honolulu, HI
		47940	Waterloo-Cedar Falls, IA*
		48300	Wenatchee, WA
		48900	Wilmington, NC
		49340	Worcester, MA-CT*
		49620	York-Hanover, PA*
		49660	Youngstown-Warren-Boardman, OH-PA*

Source: https://innovation.cms.gov/initiatives/CJR. Information for control group MSAs provided by the Centers for Medicare & Medicaid Services.

Notes: An asterisk indicates that the MSA had participating CJR hospitals in PY 6 or was included in our control group for PY 6. MSAs without an asterisk were not included in the analyses of PY 6.

CBSA = core-based statistical area; CJR = Comprehensive Care for Joint Replacement; MSA = metropolitan statistical area; PY = performance year.

2. Analytic MSA Sampling Weights and Construction of Control Group

In our analyses, we used analytic MSA-level sampling weights that reflect both the stratified sampling of MSAs used in the design of the CJR Model and the voluntary–mandatory split of the model in 2018. The probability of an MSA being selected to participate in the original CJR Model varied across the strata, with CMS proportionally under-sampling MSAs in the lower average episode payment strata (strata 1, 2, 5, and 6) and over-sampling MSAs in higher average episode payment strata (strata 3, 4, 7, and 8). **Exhibit B-3** shows the count of CJR and control group MSAs by stratum and the proportion of MSAs in each stratum that make up the CJR and control groups.

Exhibit B-3: CMS' Original Stratified Random Sample of CJR MSAs

MSA population	MSA sampling stratum	MSA average episode payment	# MSAs eligible for sampling	CJR sample		Control group sample	
				# CJR MSAs	Proportion of MSAs selected for CJR	# Control group MSAs	Proportion of MSAs in the control group
Less than median population	1	Lowest quartile	25	8	32.0%	17	68.0%
	2	2nd lowest quartile	18	6	33.3%	12	66.7%
	3	3rd lowest quartile	19	8	42.1%	11	57.9%
	4	Highest quartile	22	11	50.0%	11	50.0%
More than median population	5	Lowest quartile	15	5	33.3%	10	66.7%
	6	2nd lowest quartile	28	10	35.7%	18	64.3%
	7	3rd lowest quartile	22	9	40.9%	13	59.1%
	8	Highest quartile	22	10	45.5%	12	54.5%
		Total	171	67		104	

Source: CJR evaluation team analysis of the CMS Final Rule Medicare Program Comprehensive Care for Joint Replacement Payment Model for Acute Care Hospitals Furnishing Lower Extremity Joint Replacement Services, 80 Fed. Reg. 73273 (November 24, 2015) (codified at 42 CFR 510).

Notes: CJR = Comprehensive Care for Joint Replacement; CFR = Code of Federal Regulations; CMS = Centers for Medicare & Medicaid Services; MSA = metropolitan statistical area.

CJR was implemented as a randomized control trial, so the control group began as the MSAs that were eligible but not selected to be mandatory CJR participants. We then created MSA-specific analytic sampling weights to account for the exact details of CMS' stratified random sampling design as follows:

Step 1. We began with the 104 non-CJR MSAs.

Step 2. We applied MSA-level weights to the 104 non-CJR MSAs based on the probability that the MSA was selected into the 34 mandatory CJR MSAs through the two-step selection process.

To construct the weights in Step 2, we first calculated the probabilities of the first-stage selection for each MSA, that is, the probability that the MSA was randomly selected to be in the original set of 67 CJR MSAs. These probabilities equaled the proportion of MSAs randomly selected for CJR from each MSA sampling stratum.

Next, we calculated the probabilities of the second-stage selection, that is, the probability that the MSA was selected among the 34 mandatory CJR MSAs given that it was selected in the first stage. Those second-stage selection probabilities were more complex to calculate because the MSAs for the 34 mandatory CJR MSAs were not selected randomly, so we could not rely on simple proportions.[93] We used an average treatment effect on the treated (ATT) analysis to evaluate the impact of CJR on mandatory hospitals; thus, we used the calculated probabilities to construct MSA-level weights such that the weighted control group was representative of the CJR group. For this reason, all CJR MSAs received a weight of 1. Specifically,

[93] The MSAs were selected by ranking the original 67 CJR MSAs by historical average episode payment and retaining the top half of the sample (that is, retaining the 34 MSAs with the highest historical average episode payment).

$$Weight\ for\ mandatory\ CJR\ hospitals = 1$$

$$Weight\ for\ control\ group\ hospitals =$$

$$= \frac{(probability\ in\ treatment)}{(probability\ in\ control)} = \frac{(probability\ in\ 34\ mandatory\ CJR\ MSAs)}{(probability\ in\ 104\ non\ CJR\ MSAs)}$$

Notes: These were MSA stratum-level weights so all control group hospitals in the same MSA had the same weight.

Exhibit B-4 shows the analytic weights calculated for control group MSAs and **Exhibit B-5 shows** the full list of MSA-specific weights.

Exhibit B-4: Analytic Weights for Control Group MSAs

MSA sampling stratum	MSA	Weight
3	MSA-specific weights	Ranges from 0.00 to 0.73
4	All MSAs	1.00
7	MSA-specific weights	Ranges from 0.02 to 0.69
8	All MSAs	0.83

Source: CJR evaluation team analysis of the CMS Final Rule Medicare Program; Cancellation of Advancing Care Coordination Through Episode Payment and Cardiac Rehabilitation Incentive Payment Models; Changes to Comprehensive Care for Joint Replacement Payment Model: Extreme and Uncontrollable Circumstances Policy for the Comprehensive Care for Joint Replacement Payment Model, 82 Fed. Reg. 57066 (December 1, 2017) (codified at 42 CFR 510 and 42 CFR 512).

Notes: CJR = Comprehensive Care for Joint Replacement; CFR = Code of Federal Regulations; CMS = Centers for Medicare & Medicaid Services; MSA = metropolitan statistical area.

Exhibit B-5: Analytic Sampling Weights for Control Group MSAs

MSA sampling stratum	MSA	Weight
4	All MSAs	1.00
8	All MSAs	0.83
7	Birmingham-Hoover, AL	0.69
7	Cape Coral-Fort Myers, FL	0.69
7	Chattanooga, TN-GA	0.68
7	Columbia, SC	0.11
7	Greenville-Anderson-Mauldin, SC	0.69
3	Gulfport-Biloxi-Pascagoula, MS	0.73
3	Hattiesburg, MS	0.73
3	Huntsville, AL	0.71
3	Jonesboro, AR	0.73
7	Lexington-Fayette, KY	0.69
3	Lima, OH	0.73
7	Louisville/Jefferson County, KY-IN	0.51
3	Macon, GA	0.73
3	Manchester-Nashua, NH	0.00
7	North Port-Sarasota-Bradenton, FL	0.69
7	Philadelphia-Camden-Wilmington, PA-NJ-DE-MD	0.56
3	Punta Gorda, FL	0.73

MSA sampling stratum	MSA	Weight
7	Riverside-San Bernardino-Ontario, CA	0.02
7	Salt Lake City, UT	0.05
3	Waterloo-Cedar Falls, IA	0.73
7	Worcester, MA-CT	0.69
7	York-Hanover, PA	0.69

Source: CJR evaluation team analysis of the Medicare Program; Cancellation of Advancing Care Coordination Through Episode Payment and Cardiac Rehabilitation Incentive Payment Models; Changes to Comprehensive Care for Joint Replacement Payment Model: Extreme and Uncontrollable Circumstances Policy for the Comprehensive Care for Joint Replacement Payment Model; A Final Rule by CMS, 82 FR 57066 (December 1, 2017) (codified at 42 CFR 510 and 42 CFR 512).

Notes: CJR = Comprehensive Care for Joint Replacement, MSA = metropolitan statistical area.

3. LEJR Episode Definition

Initially, for both the CJR and control group populations, the beginning of an IP episode was triggered by an admission to a CJR participating or control group hospital (called an *anchor hospitalization*) with a resulting discharge in Medicare Severity-Diagnosis Related Group (MS-DRG) 469 or 470 (LEJR with major complications or comorbidities [MCC] and LEJR without MCC, respectively). Starting in October 2020, CMS added two new MS-DRGs for LEJR due to hip fracture (521 with MCC and 522 without MCC) as episode triggers. Previously, hip fracture episodes were discharged under MS-DRGs 469 or 470 and were identified as having a hip fracture based on International Classification of Diseases diagnosis codes. The end of the episode is 90 days after the anchor hospitalization end date.

For OP LEJRs, the beginning of the episode is triggered by an LEJR performed in the OP department of a CJR participating or control group hospital, as identified in Part B institutional claims by Current Procedural Terminology code 27447 for total knee arthroplasty or 27130 for total hip arthroplasty assigned to C-APC 5115 with status indicator "J1." The end of the episode is 90 days after the OP procedure.

Identically for both IP and OP LEJR recipients, Medicare beneficiaries who met and maintained the following eligibility throughout the period were included in the analysis:

- Enrolled in Medicare Parts A and B;
- Medicare was the primary payer; and
- Not eligible for Medicare based on end-stage renal disease (ESRD)

As specified in the Final Rule, episodes were canceled in the CJR Model and excluded from the analysis if:

- The patient no longer met the eligibility criteria described above;
- The patient was readmitted to a participating hospital during the episode and discharged under MS-DRG 469, 470, 521 or 522 (in which case the first episode is canceled and a new CJR episode begins);
- The patient died at any time during the episode period;

- The episodes started on or after July 1, 2017, and were prospectively assigned to a Next Generation ACO, an MSSP ACO in track 3, or a Comprehensive ESRD Care Model ACO with downside risk;[94] or
- The episodes were attributed to the BPCI initiative[95]

To estimate the all-cause mortality rate measure, we retained episodes that were canceled due to death of patient but otherwise met all other eligibility criteria. We also excluded episodes that lacked certain beneficiary information used to risk-adjust outcomes (for example, age, sex, and 6 months of Medicare FFS enrollment history prior to the LEJR hospital admission)

[94] This additional exclusion criterion was added with the January 2017 Final Rule, *Advancing Care Coordination Through Episode Payment Models (EPMs); Cardiac Rehabilitation Incentive Payment Model; and Changes to the Comprehensive Care for Joint Replacement Model (CJR)*. Available at: https://www.federalregister.gov/documents/2017/01/03/2016-30746/medicare-program-advancing-care-coordination-through-episode-payment-models-epms-cardiac

[95] Episodes initiated at CJR participant hospitals could be attributed to a physician group practice participating in the BPCI initiative or to skilled nursing facilities (SNFs), inpatient rehabilitation facilities (IRFs), long-term care hospitals (LTCHs), or home health agencies (HHAs) participating in the BPCI Initiative Model 3.

C. Secondary Data Claims-Based Outcomes

1. Measures of Impact on Payments, Utilization, and Quality

In this section, we present the episode-level outcome measures that we constructed using Medicare FFS claims to assess the impact of the CJR Model on Medicare payments, utilization, and quality (**Exhibit B-6**).

Exhibit B-6: Claims-Based Outcome Definitions[96]

Domain	Outcome name	Definition	Measurement period(s)	Eligible sample
Medicare payments	Total Medicare standardized allowed amounts per episode[97]	The sum of Medicare payment and beneficiary out-of-pocket amounts for related items and services covered by Medicare Part A and Part B[98] performed during the LEJR hospitalization (anchor hospitalization) through the 90-day post-discharge period that are included in the episode.	Anchor hospitalization through 90-day post-discharge period	Beneficiaries who: 1) have a complete FFS enrollment history 6 months prior to the anchor hospitalization; 2) have consistent, reliable sex and age data (age <115); 3) maintain Parts A and B enrollment throughout the measurement period; 4) have a measurement period that ends on or before December 31, 2022; 5) have non-missing Medicare standardized allowed payment information for the episode; 6) have not died from any cause during the anchor hospitalization or 90-day post-discharge period.
Medicare payments	Medicare standardized allowed amount for the anchor hospitalization per episode	The sum of Medicare payment and beneficiary out-of-pocket amounts for the LEJR anchor hospitalization (Medicare Severity-Diagnosis Related Group 469, 470, 521, or 522 for inpatient episodes covered under Medicare Part A; CPT 27447 for outpatient TKA episodes and CPT 27130 for outpatient THA covered under Medicare Part B).	Anchor hospitalization	Beneficiaries who: 1) have a complete FFS enrollment history 6 months prior to the anchor hospitalization; 2) have consistent, reliable sex and age data (age <115); 3) maintain Parts A and B enrollment throughout the measurement period; 4) have a measurement period that ends on or before December 31, 2022; 5) have non-missing Medicare standardized allowed payment information for the episode; 6) have not died from any cause during the anchor hospitalization or 90-day post-discharge period.

[96] The eligible sample column notes the inclusion criteria for episodes as defined by the Final Rule and additional measure-specific inclusion criteria required for the evaluation.

[97] Standardized payments remove wage adjustments and other Medicare payment adjustments (for example, GME, IME, and DSH). Allowed amounts include beneficiary cost sharing.

[98] Episode-related items and services paid under Medicare Part A or Part B, after exclusions are applied, including physician services, inpatient hospital services (including readmissions with certain exceptions discussed in the Final Rule), inpatient psychiatric facility services, long-term care hospital services, IRF services, SNF services, HHA services, hospital outpatient services, outpatient therapy services, clinical laboratory services, DME, Part B drugs, and hospice.

Domain	Outcome name	Definition	Measurement period(s)	Eligible sample
Medicare payments **Medicare payments (continued)**	Medicare Part A standardized allowed amounts per episode, by service	The sum of Medicare payment and beneficiary out-of-pocket amounts for readmissions, IRF, and SNF services covered under Medicare Part A. Includes all costs incurred during the 90 days following discharge.	90-day post-discharge period	Beneficiaries who: 1) have a complete FFS enrollment history 6 months prior to the anchor hospitalization; 2) have consistent, reliable sex and age data (age <115); 3) maintain Parts A and B enrollment throughout the measurement period; 4) have a measurement period that ends on or before December 31, 2022; 5) have non-missing Medicare standardized allowed payment information for the episode; 6) have not died from any cause during the anchor hospitalization or 90-day post-discharge period.
	Medicare standardized allowed amounts for HHA services per episode	The sum of Medicare payment and beneficiary out-of-pocket amounts for HHA services covered under Medicare Part A or Part B HHA.	90-day post-discharge period	Beneficiaries who: 1) have a complete FFS enrollment history 6 months prior to the anchor hospitalization; 2) have consistent, reliable sex and age data (age <115); 3) maintain Parts A and B enrollment throughout the measurement period; 4) have a measurement period that ends on or before December 31, 2022; 5) have non-missing Medicare standardized allowed payment information for the episode; 6) have not died from any cause during the anchor hospitalization or 90-day post-discharge period.
	Medicare Part B standardized allowed amounts per episode	The sum of Medicare payment and beneficiary out-of-pocket amounts for related items and services covered under Medicare Part B (except HHA services) including physician evaluation and management services, outpatient therapy services (speech, occupation, and physical therapy), imaging and lab services, procedures, DME, all other non-institutional services, and other institutional services.	90-day post-discharge period	Beneficiaries who: 1) have a complete FFS enrollment history 6 months prior to the anchor hospitalization; 2) have consistent, reliable sex and age data (age <115); 3) maintain Parts A and B enrollment throughout the measurement period; 4) have a measurement period that ends on or before December 31, 2022; 5) have non-missing Medicare standardized allowed payment information for the episode; 6) have not died from any cause during the anchor hospitalization or 90-day post-discharge period.
	Medicare standardized allowed amounts for services provided in the 30 days post-episode per episode	The sum of Medicare payment and beneficiary out-of-pocket amounts for all health care services covered under Medicare Part A or B performed during the 30-day post-episode period	30-day post-episode period	Beneficiaries who: 1) have a complete FFS enrollment history 6 months prior to the anchor hospitalization; 2) have consistent, reliable sex and age data (age <115); 3) maintain Parts A and B enrollment throughout the measurement period; 4) have a measurement period that ends on or before January 30, 2023; 5) have non-missing Medicare standardized allowed payment information for the episode; 6) have not died from any cause during the anchor hospitalization or 90-day post-discharge period.

Domain	Outcome name	Definition	Measurement period(s)	Eligible sample
Utilization	First discharge to IRF	The percentage of all episodes with beneficiaries initially discharged to an IRF. The first PAC setting is an IRF (a freestanding facility or a distinct unit within an acute hospital) if admission to the IRF occurred within the first 5 days of hospital discharge and no other PAC use occurred prior to IRF admission. If the beneficiary is directly transferred to another ACH after the anchor hospitalization, then the first PAC setting was defined within 5 days of the transfer discharge.	1st to 5th day after discharge from the anchor or transfer hospitalization	Beneficiaries who: 1) have a complete FFS enrollment history 6 months prior to the anchor hospitalization; 2) have consistent, reliable sex and age data (age <115); 3) maintain Parts A and B enrollment throughout the measurement period; 4) have a measurement period that ends on or before December 31, 2022; 5) have non-missing Medicare standardized allowed payment information for the episode; 6) have not died from any cause during the anchor hospitalization or 90-day post-discharge period.
	First discharge to SNF	The percentage of all episodes with beneficiaries initially discharged to a SNF. The first PAC setting is a SNF if admission to the SNF occurred within the first 5 days of hospital discharge and no other PAC use occurred prior to SNF admission. If the beneficiary is directly transferred to another ACH after the anchor hospitalization, then the first PAC setting was defined within 5 days of the transfer discharge.	1st to 5th day after discharge from the anchor or transfer hospitalization	Beneficiaries who: 1) have a complete FFS enrollment history 6 months prior to the anchor hospitalization; 2) have consistent, reliable sex and age data (age <115); 3) maintain Parts A and B enrollment throughout the measurement period; 4) have a measurement period that ends on or before December 31, 2022; 5) have non-missing Medicare standardized allowed payment information for the episode; 6) have not died from any cause during the anchor hospitalization or 90-day post-discharge period.
	First discharge to HHA	The percentage of all episodes with beneficiaries initially discharged to a HHA. The first PAC setting is a HHA if admission to the HHA occurred within 14 days of hospital discharge and no other PAC use occurred prior to HHA admission. If the beneficiary is directly transferred to another ACH after the anchor hospitalization, then the first PAC setting was defined within 14 days of the transfer discharge.	1st to 14th day after discharge from the anchor or transfer hospitalization	Beneficiaries who: 1) have a complete FFS enrollment history 6 months prior to the anchor hospitalization; 2) have consistent, reliable sex and age data (age <115); 3) maintain Parts A and B enrollment throughout the measurement period; 4) have a measurement period that ends on or before December 31, 2022; 5) have non-missing Medicare standardized allowed payment information for the episode; 6) have not died from any cause during the anchor hospitalization or 90-day post-discharge period.

Domain	Outcome name	Definition	Measurement period(s)	Eligible sample
Utilization (continued)	First discharge to home without HHA	The percent of all episodes with beneficiaries initially discharged to home without HHA services. The first PAC setting is home without HHA if the beneficiary is not admitted to a SNF or IRF within 5 days of hospital discharge and is not admitted to a HHA within 14 days of hospital discharge. If the beneficiary is directly transferred to another ACH after the anchor hospitalization, then the first PAC setting was defined within 14 days of the transfer discharge.	1st to 14th day after discharge from the anchor or transfer hospitalization	Beneficiaries who: 1) have a complete FFS enrollment history 6 months prior to the anchor hospitalization; 2) have consistent, reliable sex and age data (age <115); 3) maintain Parts A and B enrollment throughout the measurement period; 4) have a measurement period that ends on or before December 31, 2022; 5) have non-missing Medicare standardized allowed payment information for the episode; 6) have not died from any cause during the anchor hospitalization or 90-day post-discharge period.
	Any HHA use	The percent of all episodes with beneficiaries using any HHA services during the 90-day post-discharge period, as indicated by non-zero Medicare payment and beneficiary out-of-pocket amounts for HHA services covered under Medicare Part A or Part B.	90-day post-discharge period	Beneficiaries who: 1) have a complete FFS enrollment history 6 months prior to the anchor hospitalization; 2) have consistent, reliable sex and age data (age <115); 3) maintain Parts A and B enrollment throughout the measurement period; 4) have a measurement period that ends on or before December 31, 2022; 5) have non-missing Medicare standardized allowed payment information for the episode; 6) have not died from any cause during the anchor hospitalization or 90-day post-discharge period.
	Number of IRF days	The average number of IRF days of care during the 90-day post-discharge period.	90-day post-discharge period	Beneficiaries who: 1) have a complete FFS enrollment history 6 months prior to the anchor hospitalization; 2) have consistent, reliable sex and age data (age <115); 3) maintain Parts A and B enrollment throughout the measurement period; 4) have a measurement period that ends on or before December 31, 2022; 5) have at least one IRF day during this period; 6) have non-missing Medicare standardized allowed payment information for the episode; 7) have not died from any cause during the anchor hospitalization or 90-day post-discharge period.

Domain	Outcome name	Definition	Measurement period(s)	Eligible sample
Utilization (continued)	Number of SNF days	The average number of SNF days of care during the 90-day post-discharge period.	90-day post-discharge period	Beneficiaries who: 1) have a complete FFS enrollment history 6 months prior to the anchor hospitalization; 2) have consistent, reliable sex and age data (age <115); 3) maintain Parts A and B enrollment throughout the measurement period; 4) have a measurement period that ends on or before December 31, 2022; 5) have at least one SNF day during this period; 6) have non-missing Medicare standardized allowed payment information for the episode; 7) have not died from any cause during the anchor hospitalization or 90-day post-discharge period.
	Number of HHA visits	The average number of HHA visits during the 90-day post-discharge period.	90-day post-discharge period	Beneficiaries who: 1) have a complete FFS enrollment history 6 months prior to the anchor hospitalization; 2) have consistent, reliable sex and age data (age <115); 3) maintain Parts A and B enrollment throughout the measurement period; 4) have a measurement period that ends on or before December 31, 2022; 5) have at least one HHA visit during this period; 6) have non-missing Medicare standardized allowed payment information for the episode; 7) have not died from any cause during the anchor hospitalization or 90-day post-discharge period.
	Number of outpatient PT or OT visits	The average number of outpatient PT or OT visits during the 90-day post-discharge period.	90-day post-discharge period	Beneficiaries who: 1) have a complete FFS enrollment history 6 months prior to the anchor hospitalization; 2) have consistent, reliable sex and age data (age <115); 3) maintain Parts A and B enrollment throughout the measurement period; 4) have a measurement period that ends on or before December 31, 2022; 5) have at least one outpatient PT or OT visit during this period; 6) have non-missing Medicare standardized allowed payment information for the episode; 7) have not died from any cause during the anchor hospitalization or 90-day post-discharge period.

Domain	Outcome name	Definition	Measurement period(s)	Eligible sample
Quality	Unplanned readmission rate	The proportion of episodes with one or more unplanned readmissions for any eligible condition.[99] Following these specifications, we excluded planned admissions, based on Agency for Healthcare Research and Quality Clinical Classification System Procedure and Diagnoses codes.	90-day post-discharge period	Beneficiaries who: 1) have a complete FFS enrollment history 6 months prior to the anchor hospitalization; 2) have consistent, reliable sex and age data (age <115); 3) maintain Parts A and B enrollment throughout the measurement period; 4) have a measurement period that ends on or before December 31, 2022; 5) are discharged from the anchor hospitalization in accordance with medical advice; 6) have non-missing Medicare standardized allowed payment information for the episode ; 7) have not died from any cause during the anchor hospitalization or 90-day post-discharge period.
	ED visit rate	The proportion of episodes with one or more ED visits during the 90-day post-discharge period for which the beneficiary required medical treatment but was not admitted to the hospital. Eligible ED visits are outpatient claims with a code indicating the beneficiary used the ED but was not admitted to the hospital.	90-day post-discharge period	Beneficiaries who: 1) have a complete FFS enrollment history 6 months prior to the anchor hospitalization; 2) have consistent, reliable sex and age data (age <115); 3) maintain Parts A and B enrollment throughout the measurement period; 4) have a measurement period that ends on or before December 31, 2022; 5) are discharged from the anchor hospitalization in accordance with medical advice; 6) have non-missing Medicare standardized allowed payment information for the episode; 7) have not died from any cause during the anchor hospitalization or 90-day post-discharge period.

[99] Updated specification documents were released by CMS in March 2019 for the unplanned readmission measure, and the measure was revised accordingly. Available at: https://www.qualitynet.org/inpatient/measures/readmission/methodology

Domain	Outcome name	Definition	Measurement period(s)	Eligible sample
Quality (continued)	All-cause mortality rate	Death from any cause during the anchor hospitalization or 90-day post-discharge period.	Anchor hospitalization and 90-day post-discharge period	Under the CJR Model, death during the anchor hospitalization or 90-day PDP cancels the episode. Therefore, this analysis includes CJR and control group episodes as well as beneficiaries at CJR participant and control group hospitals that would have been identified as episodes if they had not died during the episode of care. Beneficiaries who: 1) have a complete FFS enrollment history 6 months prior to the anchor hospitalization; 2) have consistent, reliable sex and age data (age <115); 3) maintain Parts A and B enrollment throughout the measurement period; 4) have not received hospice care in the 6 months prior to admission; 5) have a measurement period that ends on or before December 31, 2022; 6) are discharged from the anchor hospitalization in accordance with medical advice; 7) have non-missing Medicare standardized allowed payment information for the episode.
	Incidence of any complications	The proportion of elective episodes with incidence (during the anchor hospitalization or a readmission) of: acute myocardial infarction, pneumonia, or sepsis or septicemia within the 7-day PDP; or surgical site bleeding or pulmonary embolism within the 30-day PDP; or mechanical complications, periprosthetic joint infection, or wound infection within the 90-day PDP. This measure was based on specifications for the NQF-endorsed THA or TKA complications measure (NQF measure 1550).[100] Death in the 30 days after discharge is part of the technical definition, but is not included in our analysis because beneficiaries who died during the anchor hospitalization or in the 90-day PDP are excluded from the CJR Model.	90-day post-discharge period	Beneficiaries who: 1) have a complete FFS enrollment history 6 months prior to the anchor hospitalization; 2) have consistent, reliable sex and age data (age <115); 3) maintain Parts A and B enrollment throughout the measurement period; 4) have a measurement period that ends on or before December 31, 2022; 5) are discharged from the anchor hospitalization in accordance with medical advice; 6) have non-missing Medicare standardized allowed payment information for the episode; 7) have not died from any cause during the anchor hospitalization or 90-day post-discharge period.

[100] Updated specification documents were released by CMS in March 2019 for the THA or TKA complications measure, and the measure was revised accordingly. Available at: https://www.qualitynet.org/inpatient/measures/complication/methodology

Note: ACH = acute care hospital, CJR = Comprehensive Care for Joint Replacement, CMS = Centers for Medicare & Medicaid Services, CPT = current procedural terminology, DME = durable medical equipment, DSH = disproportionate share hospital, ED = emergency department, FFS = fee-for-service, GME = graduate medical education, HHA = home health agency, IME = indirect medical education, IRF = inpatient rehabilitation facility, LEJR = lower extremity joint replacement, NQF = National Quality Forum, OT = occupational therapy, PAC = post-acute care, PDP = post-discharge period, PT = physical therapy, SNF = skilled nursing facility, THA = total hip arthroplasty, TKA = total knee arthroplasty.

2. Measures of Unintended Consequences

In our evaluation of unintended consequences of the CJR Model, one of the topics we studied was changes in patient mix.[101] **Exhibit B-7** lists the patient characteristics from claims and enrollment data that we monitored. Although the impact analysis on payment, utilization, and quality controlled for changes in many of these patient characteristics, we also monitored changes in these characteristics separately to directly examine changes in patient mix.

Exhibit B-7: Measures of Patient Mix

Type of unintended consequence	Measure name or description
Changes in patient mix	Age • Less than 65 years • 65-74 years • 75-84 years • 85 years or greater
	Sex
	Race and ethnicity
	Medicaid eligibility
	Disability, no ESRD
	HCC • Score • Count • Indicator for having a count of 0, 1, 2, 3, or 4 or more
	Congestive heart failure
	Dementia
	Obesity
	Hypertension
	Diabetes
	Prior utilization (in the 6 months prior to the anchor hospitalization) • IP ACH stay • IRF stay • SNF stay • HH use • Any prior care[a]

Source: Patient mix measures were constructed from Medicare fee-for-service claims and beneficiary enrollment data.

Notes: ACH = acute care hospital; ESRD = end-stage renal disease; HCC = Hierarchical Condition Category; HH = home health; IRF = inpatient rehabilitation facility; IP = inpatient; SNF = skilled nursing facility.

[a] Any prior care included IP hospital, psychiatric hospital, emergency department, SNF, IRF, home health, long-term care hospital, and hospice during the 6 months prior to anchor hospitalization.

[101] As presented in **Chapter VI: Analysis of Potential Unintended Consequences of the CJR Model**, we also completed analyses on the model's impact on payments in the 30 days following the episode and analyses on the model's impact on the total market volume of elective LEJR discharges.

We also looked at the impact of the CJR Model on payments in the 30 days following the episode in our evaluation of unintended consequences. This claims-based measure is defined in detail in **Exhibit B-6**.

D. Analytic Methodology

To control for both observed and unobserved differences and to isolate the impact of the CJR Model on outcomes, we employed a difference-in-differences (DiD) regression approach using the control group created by the model's randomization of CJR MSAs and supplemented by risk adjustment.

1. Difference-in-Differences Estimator

The DiD approach quantified the impact of the CJR Model by comparing changes in outcomes between the baseline and the intervention period of interest (PY 6) for the CJR population and the control group population. One of the main advantages of this approach is that it can successfully isolate the effect of unobserved characteristics of treatment and control groups that are time invariant.[102]

a. Baseline Period

The baseline period for our evaluation encompassed episodes that started between January 1, 2012, and December 31, 2014, and ended between April 1, 2012, and March 31, 2015.

b. Intervention Period

The intervention period for this annual report followed the model timeline for PY 6 and was specified as episodes ending between October 1, 2021, and December 31, 2022.[103]

c. Exclusion of Performance Years 1–5

In this annual report, we were focused on studying the impact of the CJR Model during PY 6. Thus, the methodology did not need to account for artifacts of the CJR Model prior to PY 6 if they were no longer relevant in PY 6. For example, we no longer needed to account for differential rates of OP procedures at CJR and control hospitals during PY 3–5, as we did in the Fifth Annual Report.

We conducted empirical testing both including and excluding PY 1–5 in our DiD regressions that estimate an impact for PY 6. We determined that including earlier PYs in the sample did little to improve the precision of the risk-adjusters or meaningfully change the primary estimates of interest, so we excluded the earlier PYs from our impact analyses of PY 6.

[102] Although the DiD model controls for unobserved heterogeneity that is fixed over time, it does not control for unobserved heterogeneity that varies over time.

[103] Final Rule issued by CMS in the Federal Register on April 29, 2021, extending the CJR Model: https://www.federalregister.gov/documents/2020/02/24/2020-03434/medicare-program-comprehensive-care-for-joint-replacement-model-three-year-extension-and-changes-to.

d. Regression Specification

The DiD model used an outcome measure, *Y*, and estimated the differential change in *Y* for beneficiaries receiving care from CJR participant hospitals between the baseline and the intervention period relative to that same change for beneficiaries receiving care from hospitals in the control group.

To illustrate the DiD approach, we defined:

- $Y_{i,k,t}$ as the outcome for the i^{th} episode with an LEJR at hospital *k* in period *t* (*t* = 1 during the CJR PY 6 intervention quarters and 0 otherwise)
- $CJR_{i,k}$ as an indicator that takes the value of 1 if the i^{th} episode was initiated by a CJR participant hospital *k* and takes the value of 0 otherwise.
- $X_{i,k,t}$ as hospital, geographic, and patient characteristics in period *t.*
- E[Y|t, CJR, X] as the expected value of outcome measure Y conditional on values of *t*, *CJR*, and *X.*

The DiD estimator is then:

$$DiD = [E(Y \mid t = 1, CJR = 1, X) - (E(Y \mid t = 0, CJR = 1, X)] - [E(Y \mid t = 1, CJR = 0, X) - (E(Y \mid t = 0, CJR = 0, X)] \quad \textbf{(1)}$$

To illustrate the calculation of the DiD, consider the linear model listed below:

$$Y_{i,k,t} = b_0 + b_1 \cdot t + b_2 \cdot CJR_{i,k} + b_3 \cdot CJR_{i,k} \cdot t + X_{i,k,t}{}' \cdot B + u_{i,k,t} \quad \textbf{(2)}$$

- The value of coefficient b_1 captured aggregate factors that could cause changes in outcome *Y* in the intervention period relative to the baseline period that are common across CJR and control group episodes.
- Coefficient b_2 captured the relative differences in outcomes between CJR and control group episodes.
- Coefficient b_3 determined the differential in outcome *Y* experienced by beneficiaries receiving services from CJR hospitals during the CJR intervention period relative to control group episodes in the intervention period and represented the DiD estimator.
- The vector of coefficients *B* measured the differential effects of risk factors (*X*) on the outcome variable.

Finally, to calculate the DiD estimate for outcome measures that were risk-adjusted with nonlinear models, we used the regression model's coefficient estimates to calculate each of the four conditional expectations that make up the DiD estimator in Equation 1. For all DiD models, we assessed statistical significance at the 10% level. Given the design of the CJR Model, which involved randomly sampling MSAs to participate, we accounted for clustering at the MSA level in the estimation of our standard errors in all regression models.

Additionally, we conducted various tests of whether the CJR and control populations were on parallel trends in the baseline period and sensitivity analyses related to COVID-19. See **Appendix D: Claims-Based Impact Analyses – Detailed Tables** for a description of the methodologies and results.

e. *Risk Adjustment to Control for Differences in Beneficiary Demographics and Clinical Risk Factors*

In the DiD models, we controlled for potential differences in beneficiary demographics, clinical characteristics observed before hospitalization, and provider characteristics (represented by $X_{i,i,t}$ in Equation 2 above). Demographic factors included age categories, sex, age and sex interactions, race and ethnicity indicators, Medicaid eligibility status, and disability status. We risk-adjusted all outcomes for the episode's hip fracture status, procedure type (hip or knee), and presence or absence of an MCC, defined by MS-DRGs.[104] To control for participation in other Medicare initiatives, we used an indicator variable that indicated whether the beneficiary was in the MSSP ACO, Pioneer ACO Model, or Next Generation ACO Model during the episode.[105] To control for prior health conditions, we used Hierarchical Condition Category (HCC) indicators for the 12 months preceding the anchor hospitalization,[106] as well as indicators for obesity, diabetes, hypertension, and tobacco use, generated from the claims data. To further control for case-mix differences, we included measures of the following types of prior care use: acute care IPPS hospital, skilled nursing facility (SNF), inpatient rehabilitation facility (IRF), hospice, other Part A IP, home health agency (HHA), custodial nursing facility, and others. We used the Long-Term Care MDS 3.0 data to create a measure of prior custodial nursing facility use in the 6 months prior to the episode. The collection of this assessment data was temporarily paused during most of the COVID-19 Public Health Emergency; thus, we imputed seemingly missing values of this covariate during that time period. We also included COVID-19-related risk-adjusters. To address beneficiary-level effects of the COVID-19 pandemic, we controlled for a COVID-19 diagnosis in the 30 days prior to anchor hospitalization or during the anchor hospitalization from claims data (confirmed positive, suspected, or probable with symptoms or exposure).[107]

We also controlled for provider characteristics that might have been related to the outcomes of interest, such as hospital bed count, for-profit status, previous BPCI initiative LEJR experience, and previous BPCI initiative experience in a clinical episode other than LEJR. In October 2018, the BPCI Advanced initiative began. This Innovation Center 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. To account for contamination in our control group by this other Innovation Center model, we included indicator variables that identify control group LEJR episodes performed by surgeons or at hospitals participating in the BPCI Advanced model.

---

[104] We made two modifications to DRGs for this process: (1) We back-code DRGs that occurred prior to the separation of elective and fracture DRGs to match what they would have been under this change, and (2) we code all OP procedures as MS-DRG 470.

[105] Beneficiaries with episodes during or after July 2017 that were aligned with the MSSP track 3, Next Generation ACO, or the Comprehensive ESRD Care model were excluded from the CJR Model.

[106] The CMS-HCC model is a prospective risk-adjustment model used by CMS to adjust Medicare Part C capitation payments for beneficiary health spending risk. The model adjusts for demographic and clinical characteristics. The clinical component of the model uses diagnoses from qualifying services grouped into numerous HCC indicators.

[107] One notable difference from the Fifth Annual Report is that in our main impact analysis, we did not control for county-level COVID-19 controls. The end of the Public Health Emergency during PY 6 meant that too few counties were providing reliable data for us to use at our level of desired granularity. After examining several options, we decided to drop county-level COVID-19 controls from our set of risk-adjusters for these analyses.

See **Exhibit B-8** for additional details about our risk-adjustment methodology.[108]

Exhibit B-8: Predictive Risk Factors Used to Risk-Adjust Claims-Based Outcomes for Impact Analyses

Domain	Variables
Characteristics of the procedure	• Anchor Medicare Severity-Diagnosis Related Group • Hip fracture status • Procedure type (hip or knee)
Patient demographics and enrollment	• Age (under 65, 65-79, 80+) • Sex • Race and ethnicity • Medicaid status • Disability status at enrollment in Medicare (not end-stage renal disease) • Attribution to Medicare Shared Savings Program, Pioneer ACO Model, or Next Generation ACO Models during the Comprehensive Care for Joint Replacement episode
Prior health conditions	• CMS-HCC version 21 indicators from qualifying services and diagnoses from claims and data for 12 months preceding the anchor hospitalization [a] • Obesity indicator • Diabetes indicator • Hypertension indicator • Tobacco use indicator
Prior use	• Prior care use (any acute care IP, emergency department visits, IRF, SNF, HHA, hospice, other Part A IP, long-term care hospital, and custodial nursing facility service) variables used in risk adjustment varied by model[b] ▪ Binary indicators for any care use in the 6 months preceding the start of the episode ▪ Binary indicators for any care use in the 1 month preceding the start of the episode ▪ Number of days of care use in the 6 months preceding the start of the episode
Geography	• State indicators
COVID-19	• Beneficiary-level COVID-19 diagnosis in the 30 days prior to the anchor hospitalization • Beneficiary-level COVID-19 diagnosis during the anchor hospitalization
Hospital provider characteristics	• Bed count • For-profit status • BPCI LEJR experience • BPCI experience in a clinical episode other than LEJR • LEJR performed by surgeons or at hospitals participating in the BPCI-A model for LEJR clinical episodes (control group only)

Source: Risk adjustment variables were contructed from Medicare fee-for-service claims and beneficiary enrollment data, December 2016 provider of services, fiscal year 2016 CMS Annual Inpatient Prospective Payment System, CMS Master Data Management, BPCI initiative participant list, and BPCI-A initiative participant list.

Notes: ACO = Accountable Care Organization; BPCI = Bundled Payments for Care Improvement; BPCI-A = Bundled Payments for Care Improvement Advanced; CMS = Centers for Medicare and Medicaid Services Hierarchical Condition Categories; COVID-19 = coronavirus disease 2019; HCC = hierarchical condition category; HHA = home health agency;

[108] For additional information on how we chose these risk-adjustment covariates, please see the CJR Model evaluation's First Annual Report appendices: https://www.cms.gov/files/document/cjr-firstannrptapppdf.pdf

IP = inpatient; IRF = inpatient rehabilitation facility; LEJR = lower extremity joint replacement; SNF = skilled nursing facility.

[a] The specification for each prior use variable varied for each outcome. The binary 6-month indicators were used for: SNF payment, IRF payment, HHA payment, Part B payment, unplanned readmissions, ED use, number of SNF days, and first discharge setting. The binary 1-month indicators were used for: complications and mortality. The indicators for number of days in the past 6 months were used for: total episode payment, readmissions payment, 30-day post-episode payment, number of IRF days, number of HHA visits, any HHA visits, and number of outpatient PT or OT visits.

[b] The HCC indicators in the risk adjustment model included: sepsis, different types of cancer, diabetes, obesity, malnutrition, rheumatoid arthritis, coagulation defects, dementia, drug or alcohol dependence, mood disorder, Parkinson's disease, seizure disorders, cardio-respiratory failure, congestive heart failure, angina, heart arrhythmias, stroke, vascular disease, chronic obstructive pulmonary disease, macular degeneration, kidney disease, and renal failure. (Pope, Gregory C.; Kautter, John; Ellis, Randall P.; Ash, Arlene S.; Ayanian, John Z.; Iezzoni, Lisa I.; Ingber, Melvin J.; Levy, Jesse M.; and Robst, John, "Risk adjustment of Medicare capitation payments using the CMS-HCC model" (2004). Quantitative Health Sciences Publications and Presentations. Paper 723.)

2. Model Types

We used a variety of models including logistic, Poisson, multinomial logit, ordinary least squares (OLS) regressions, and two-part models (**Exhibit B-9**). Models were estimated depending on the type and characteristics of the outcome measure. For example, logistic models were estimated for the discrete quality outcomes (that is, all claims-based quality-of-care measures). A multinomial logit model was applied to the first discharge setting. OLS models were estimated for the continuous total number of days or visits measures as well as total episode payments and Part B payments. We used two-part models for payment outcomes where a considerable number of individuals had zero payments for the particular outcome.

Exhibit B-9: Outcomes by Model Type

Model type	Outcomes	
OLS	• Total episode payments • Part B payments • Number of IRF days	• Number of SNF days • Number of HHA visits • Number of OP PT or OT Visits
Two-part models (Probit or OLS)	• Readmission payments • IRF payments • SNF payments	• HHA payments • 30-day post-episode payments
Multinomial logistic	• First post-acute discharge was to IRF • First post-acute discharge was to SNF	• First post-acute discharge was to HHA • Discharge to home without HH
Logistic	• Any HHA visits • Unplanned readmission • ED visit	• Complications • All-cause mortality

Notes: ED = emergency department; HH = home health; HHA = home health agency; IRF = inpatient rehabilitation facility; OLS = ordinary least squares; OP = outpatient; OT = occupational therapy; PT = physical therapy; SNF = skilled nursing facility.

We used estimates from the multivariate regression models to construct model-predicted outcomes (sometimes labeled "risk-adjusted mean outcomes") under two scenarios (baseline and PY 6 intervention) for both CJR and control group hospitals. To control for changes in service and case mix over time, as well as differences between CJR and non-CJR beneficiaries, we used the same reference population of beneficiaries to calculate predicted outcomes for CJR and control group episodes. The reference population used for this report was all CJR beneficiaries during the baseline and intervention period.

E. Savings to Medicare due to the CJR Model

We calculated Medicare savings by subtracting reconciliation payments to CJR participant hospitals from the change in nonstandardized paid amounts due to the CJR Model. Medicare savings were calculated on both a total and a per-episode basis.

Medicare savings = Change in nonstandardized paid amounts – Reconciliation payments

1. Change in Nonstandardized Paid Amounts

To best capture the actual amount of payments sent and received by CMS, we use nonstandardized paid amounts for our analyses on the savings to Medicare. We calculate the change in nonstandardized paid amounts using estimates from a DiD model of per-episode standardized paid amounts. The DiD estimates were multiplied by –1 and converted to nonstandardized paid amounts using a ratio of nonstandardized-to-standardized Medicare paid amounts from CJR intervention episodes **(Exhibit B-10)**. This method produced a per-episode estimate of the change in nonstandardized paid amounts. We calculated the total change in nonstandardized paid amounts by multiplying the per-episode estimate by the total number of PY 6 CJR episodes.

Exhibit B-10: Ratios of Nonstandardized-to-Standardized Medicare Paid Amounts Over Time

Time period	Mandatory hospitals
Baseline	1.038
PY 6	1.059

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 that ended between October 1, 2021, and December 31, 2022 (PY 6 intervention).

Notes: The ratio was calculated as the average nonstandardized (actual) paid amounts divided by the average standardized paid amounts for episodes. The anchor payment (Medicare Severity-Diagnosis Related Group payment for inpatient episodes) was subtracted from the total episode payment before calculating the ratio.
CJR = Comprehensive Care for Joint Replacement;
PY = performance year.

Note that in this report, we estimated Medicare savings only for PY 6. We considered the Medicare savings results for PY 1–5 reported in the Fifth Annual Report to be authoritative in terms of describing Medicare savings over the first 5 PYs of the model.[109]

2. Reconciliation Payments

We defined reconciliation payments as the total payments made to CJR participants by Medicare minus repayments from CJR participants to Medicare. Reconciliation payments could be positive or negative. In the program literature, they are often referred to as *net payment reconciliation amounts* (NPRA). The CMS CJR payment contractor provided this data. We calculated

[109] Centers for Medicare & Medicaid Services. (2023). *Comprehensive Care for Joint Replacement Model - Fifth Annual Report*. https://www.cms.gov/priorities/innovation/data-and-reports/2023/cjr-py5-annual-report

reconciliation payments per episode by dividing total reconciliation payments by the total number of CJR episodes.

In analyzing the distribution of NPRA to hospitals, we ordered hospitals by their NPRA in PY 6 and grouped the ordered hospitals into deciles, each comprising 10% of the total hospital participants. For each decile, we calculated the total NPRA received by hospitals in the decile and average NPRA across hospitals in the decile. We previously performed a similar analysis for total NPRA in PY 1–5 in the Fifth Annual Report.[110] We then conducted exploratory analyses aimed at the characteristics of hospitals that received the most reconciliation payments or had the largest repayments.

3. Considerations

In the estimation of Medicare program savings, we dropped some episodes that resulted in reconciliation payments or repayments to CMS from the estimation sample, primarily due to our requirement that beneficiaries have a complete FFS enrollment history for 6 months prior to the anchor hospitalization.[111] We did not extrapolate estimated payment reductions to these "missing" episodes. Thus, our estimates of total Medicare program savings are slightly conservative; if the missing episodes also had payment reductions due to the CJR Model, our estimates of Medicare program savings are underestimates.

F. Patient Survey

We developed the CJR patient survey to explore differences between CJR and control patients in functional status and pain, need of caregiver help, care experience, and overall satisfaction at the end of the episode. The patient-reported outcomes in the survey captured information not available from other data sources, such as claims or assessment data. Results of the patient survey can be found in **Appendix E: Patient Survey Results**.

1. Data Collection

a. Survey Sample

We surveyed a sample of LEJR patients during a 4-month window of PY 6. We administered the patient survey in two batches, with each batch covering 2 months of episodes: July–August 2022 or September–October 2022. In each batch, we took a census of beneficiaries with surgery for hip fractures from the CJR and control groups. Within the CJR and control groups, we then selected a proportional random sample of beneficiaries with elective IP and OP surgeries, stratified by four categories of age (<65, 65–74, 75–84, 85+ years) to receive the survey. Starting sample sizes were designed to target 1,300 total CJR and 1,300 total CJR responses from beneficiaries with IP episodes (including the census of hip fracture episodes). Estimated starting samples for elective episodes were based on response rates from Waves 5 and 6 of the survey (2019) and a power analysis with a target power of 0.80, a significance level of 5%, and an impact equal to 3% of prehospital functional status measures. Because OP episodes were not part of our sample in

[110] Centers for Medicare & Medicaid Services. (2023). *Comprehensive Care for Joint Replacement Model - Fifth Annual Report*. https://www.cms.gov/priorities/innovation/data-and-reports/2023/cjr-py5-annual-report

[111] See the Fourth Annual Report appendices for information and results of sensitivity analyses related to these episodes: https://www.cms.gov/priorities/innovation/data-and-reports/2022/cjr-py4-ar-app

Waves 5 and 6 (as they were not yet incorporated into CJR), we used the response rate of elective IP episodes with MS-DRG 470.

In selecting control beneficiaries into our starting sample, we exact-matched control beneficiaries to our selected CJR sample by age. Within each age cell, we also used nearest-neighbor propensity score matching to match control respondents by sex, race and ethnicity, dual eligibility, disability, knee versus hip procedure, MS-DRG (for IP episodes), MSA randomization stratum, hospital ownership, hospital academic status, hospital size (in beds), and whether the hospital or surgeon had participated in the LEJR bundle under BPCI Model 2. Detailed information about the episode count and survey sample for the two patient survey batches can be found in **Exhibit B-11**.

Exhibit B-11: Total LEJR Episodes and Starting Survey Sample of LEJR Patients, by Survey Batch and Surgery Type

Batch	Discharge date	Sample	CJR LEJR episodes	CJR patients sampled	Control LEJR episodes	Control patients sampled[a]
1	July or August 2022	All-LEJR	7,532	4,610	8,545	4,494
		Elective inpatient	1,928	1,627	1,702	1,540
		Elective outpatient	4,671	2,050	5,873	1,984
		Hip fracture	933	933	970	970
2	September or October 2022	All-LEJR	7,504	4,594	8,392	4,515
		Elective inpatient	1,967	1,628	1,629	1,547
		Elective outpatient	4,619	2,048	5,779	1,984
		Hip fracture	918	918	984	984

Source: CJR evaluation team analysis of survey data for patients with inpatient discharge or outpatient procedure in July – October 2022.

Notes: CJR = Comprehensive Care for Joint Replacement; LEJR = lower extremity joint replacement; MSA = metropolitan statistical area.

[a] We intentionally sampled fewer control respondents among elective episodes based on response rate data from prior survey waves that indicated higher response rates for control MSAs.

b. Survey Administration

For each of our two batches, we mailed surveys to all selected patients two months after the end of our two-month sampling window, allowing time for claims runout, sampling, and processing of sample for mailing. This meant that patients received their survey between 60 and 120 days after their lower extremity joint replacement (LEJR) inpatient discharge or outpatient procedure (97 days later, on average) depending on when in our two-month sampling window they underwent surgery. Reminder postcards were sent one week later. Four weeks after the initial mailing, we mailed non-respondents a second survey. Outbound telephone and email follow-up with non-respondents began approximately ten weeks after the first mailing. On average, respondents returned the survey 127 days after leaving the hospital and the difference between CJR and control respondents was less than one day (p=0.298). Sensitivity analysis from prior waves did not find any evidence that results varied when we controlled for time between leaving the hospital and survey receipt.

c. Response Rates and Analytic Samples

The average response rate in our sample was 53.6% across all LEJRs (**Exhibit B-12**). The response rate was lowest for hip fracture episodes (31.9%) and highest for elective OP episodes (61.6%). There were no significant differences in response rates between CJR and control respondents. In total, our sample included responses representing 288 out of 310 CJR PY 6 participants with at least one episode during our sampling period and 286 out of 308 control hospitals with at least one episode. For each of the elective IP, elective OP, and hip fracture samples, responses represented 217 out of 269 CJR hospitals, 217 out of 250 CJR hospitals, and 209 out of 279 CJR hospitals, respectively, with similar representation among control hospitals. We fell short of our targeted 2,600 IP and OP responses for the CJR and control groups; thus, our overall sample size (4,937 CJR responses and 4,740 control responses) was slightly below our targeted 5,200 CJR responses and 5,200 control responses in total. However, our sample size was large enough to obtain estimates with 95% confidence intervals comparable to previous waves.

Exhibit B-12: Sample Size and Response Rate for LEJR Patients, by Type of Surgery

Group	Patients surveyed (starting sample)		Survey responses received (analytic sample)		Response rate		
	CJR	Control	CJR	Control	CJR	Control	p-value
All-LEJR	9,204	9,009	4,937	4,740	53.6%	52.6%	0.48
Elective IP	3,255	3,087	1,820	1,674	55.9%	54.2%	0.49
Elective OP	4,098	3,968	2,526	2,400	61.6%	60.5%	0.45
Hip fracture	1,851	1,954	591	666	31.9%	34.1%	0.17

Source: CJR evaluation team analysis of survey data for patients with IP discharge or OP procedure in July–October 2022.

Notes: CJR = Comprehensive Care for Joint Replacement; IP = inpatient; LEJR = lower extremity joint replacement; OP = outpatient.

d. Patient Survey Questions

The patient survey instrument consisted of the following questions:

---Begin survey----

1. Who is completing this survey?
   - ○ Person named in the cover letter
   - ○ Person named in the cover letter, with help from a family member, friend or caregiver
   - ○ A family member, friend, or caregiver of the person named in the cover letter
   - ○ If the person to whom this was mailed cannot complete the survey, and there is no one else who can do it for him or her, please mark this response and return the blank survey

Section 1. Before the Hospital

We would like to know how you were doing ***before*** you went to the hospital listed in the cover letter to have your joint replaced.

2. Did you have any sessions with a physical therapist for the joint you had replaced in the ***two weeks or so before*** your joint replacement surgery?
   - ○ Yes
   - ○ No
   - ○ Don't know/Don't remember

The next questions ask about the ***week before*** your joint replacement surgery.

3. Thinking about the week before your joint replacement surgery, how often did pain in the joint that you had replaced interfere with your normal activities?
   - ○ All of the time
   - ○ Most of the time
   - ○ Some of the time
   - ○ A little of the time
   - ○ None of the time
   - ○ Don't know/Don't remember

4. Thinking about the week before your joint replacement surgery, were you taking any of the following types of ***medications specifically*** for pain in the joint that you had replaced?
   - ○ Prescription pain medication only
   - ○ Over the counter pain medication only
   - ○ Both prescription and over the counter pain medications
   - ○ No medication for pain in the joint that was replaced
   - ○ Don't know/Don't remember

5. Thinking about the week before your joint replacement surgery, what best describes your ***use of a mobility aid*** such as a wheelchair, scooter, walker, or cane?
   - ○ I never used a mobility aid
   - ○ I sometimes used a mobility aid
   - ○ I always used a mobility aid
   - ○ Don't know/Don't remember

6. Thinking about the week before your joint replacement surgery, what best describes your ability to ***walk by yourself without resting***? That is, walk without the help of another person or the help of a mobility aid.
   - ○ I could walk more than several blocks by myself without resting
   - ○ I could walk several blocks by myself without resting
   - ○ I could walk one block by myself without resting
   - ○ I could walk from one room to another by myself without resting
   - ○ I was not able to walk by myself without resting
   - ○ Don't know/Don't remember

7. Thinking about the week before your joint replacement surgery, how much difficulty did you have ***walking up or down 12 stairs***?
   - ○ I had no difficulty walking up or down 12 stairs
   - ○ I had some difficulty walking up or down 12 stairs
   - ○ I had a lot of difficulty walking up or down 12 stairs
   - ○ I was not able to walk up or down 12 stairs
   - ○ Don't know/Don't remember

8. Thinking about the week before your joint replacement surgery, how much ***difficulty*** did you have rising from sitting?
   - ○ Extreme
   - ○ Severe
   - ○ Moderate
   - ○ Mild
   - ○ None
   - ○ Don't know/Don't remember

9. Thinking about the week before your joint replacement surgery, how much ***difficulty*** did you have standing?
   - ○ Extreme
   - ○ Severe
   - ○ Moderate
   - ○ Mild
   - ○ None
   - ○ Don't know/Don't remember

10. Thinking about the week before your joint replacement surgery, how much ***difficulty*** did you have getting on/off the toilet?
    - ○ Extreme
    - ○ Severe
    - ○ Moderate
    - ○ Mild
    - ○ None
    - ○ Don't know/Don't remember

Section 2. After the Hospital

Now we'd like to learn about your experience ***after you left the hospital*** listed in the cover letter, and the weeks immediately after.

11. Thinking about when you left the hospital for your joint replacement surgery, would you say that you were…
    - ○ Discharged too early
    - ○ Discharged at the right time or
    - ○ Discharged too late
    - ○ Don't know/Don't remember

12. Thinking about the care you received – in the ***two weeks after*** your joint replacement surgery – from doctors, nurses and therapists, at home, in a doctor or therapist's office or in a medical facility – how would you rate the level of care overall?
    - ○ Level of care during two weeks after surgery was more than I needed
    - ○ Level of care during two weeks after surgery was about right
    - ○ Level of care during two weeks after surgery was not enough
    - ○ Don't know/Don't remember

13. Do you live in your own home, in someone else's home, or in an assisted living facility?
    - ○ Yes
    - ○ No, **Go To Section 3 on page I-5**

14. When you went home after your joint replacement surgery, did you have all the medical equipment you needed (for example, walker, elevated commode, grabber, shower chair, device to help put on socks)?
    - ○ Yes
    - ○ No
    - ○ Don't know/Don't remember

We would like to learn about the help you received from other people when you went home after your joint replacement surgery, or to someone else's home or an assisted living facility.

15. Thinking back to the people who helped you, who was your ***main caregiver***, that is, the person who helped you the most after your joint replacement surgery?
    - ○ Spouse/partner
    - ○ Adult child
    - ○ Another relative
    - ○ Paid caregiver
    - ○ Friend, neighbor, or someone else
    - ○ No help at home after joint replacement surgery

16. When you went home after joint replacement surgery, how much help did you need from your main caregiver with ***putting on or taking off your clothes***?
    - ○ No help needed
    - ○ Some help needed
    - ○ Complete help needed
    - ○ Don't know/Don't remember

17. When you went home after joint replacement surgery, how much help did you need from your main caregiver with using the ***toilet***?
    - ○ No help needed
    - ○ Some help needed
    - ○ Complete help needed
    - ○ Don't know/Don't remember

18. When you went home after joint replacement surgery, how much help did you need from your main caregiver with ***bathing***?
    - ○ No help needed
    - ○ Some help needed
    - ○ Complete help needed
    - ○ Don't know/Don't remember

Section 3. Health Care Experiences in-Hospital and After

We want to learn about your experiences while you were ***in the hospital*** listed in the cover letter ***and any other place where you received medical care*** following that hospitalization.

In the following questions, the term ***"health care providers"*** means doctors, nurses, physical or occupational therapists and any other medical professionals who helped take care of you during your time in the hospital and afterwards, in other facilities or at home in any capacity.

Please think of all these types of providers and locations when rating your ***level of satisfaction*** in the next few questions.

19. How satisfied or dissatisfied were you with the extent to which health care providers ***listened to your thoughts and preferences about your medical treatment***?
    - ○ Very dissatisfied
    - ○ Somewhat dissatisfied
    - ○ Neither satisfied nor dissatisfied
    - ○ Somewhat satisfied
    - ○ Very satisfied

20. How satisfied or dissatisfied were you with the ***place you were sent after you left the hospital***, for example, home, rehabilitation facility, nursing home, long-term care hospital?
    - ○ Very dissatisfied
    - ○ Somewhat dissatisfied
    - ○ Neither satisfied nor dissatisfied
    - ○ Somewhat satisfied
    - ○ Very satisfied

21. How satisfied or dissatisfied were you with the ***coordination of your care among doctors, nurses, and therapists*** in the hospital and after discharge?
    - ○ Very dissatisfied
    - ○ Somewhat dissatisfied
    - ○ Neither satisfied nor dissatisfied
    - ○ Somewhat satisfied
    - ○ Very satisfied
    - ○ Don't know

22. How satisfied or dissatisfied were you with the ***instructions you received from doctors, nurses, and therapists*** about your treatment?
    - ○ Very dissatisfied
    - ○ Somewhat dissatisfied
    - ○ Neither satisfied nor dissatisfied
    - ○ Somewhat satisfied
    - ○ Very satisfied

23. How satisfied or dissatisfied were you with your ***overall recovery*** from joint replacement surgery ***since you left the hospital***?
    - ○ Very dissatisfied
    - ○ Somewhat dissatisfied
    - ○ Neither satisfied nor dissatisfied
    - ○ Somewhat satisfied
    - ○ Very satisfied

Section 4. How are you Feeling Today?

24. In the past week, how much does ***pain in the joint that you had replaced*** currently interfere with your normal activities?
    - ○ All of the time
    - ○ Most of the time
    - ○ Some of the time
    - ○ A little of the time
    - ○ None of the time
    - ○ Don't know/Don't remember

25. Thinking about the past week, have you been taking any of the following types of medications ***specifically*** for pain in the joint you had replaced?
    - ○ Prescription pain medication only
    - ○ Over the counter pain medication only
    - ○ Both prescription and over the counter pain medications
    - ○ No medication for pain in the joint that was replaced
    - ○ Don't know/Don't remember

26. What best describes your ***use of a mobility aid*** over the past week, such as a wheelchair, scooter, walker or cane?
    - ○ I never use a mobility aid
    - ○ I sometimes use a mobility aid
    - ○ I always use a mobility aid
    - ○ Don't know/Don't remember

27. What best describes your current ability to ***walk by yourself without resting***? That is, without the help of another person or the help of a mobility aid?
    - ○ I can walk more than several blocks by myself without resting
    - ○ I can walk several blocks by myself without resting
    - ○ I can walk one block by myself without resting
    - ○ I can walk from one room to another by myself without resting
    - ○ I am not able to walk by myself without resting
    - ○ Don't know/Don't remember

28. How much difficult do you currently have ***walking up or down 12 stairs***?
    - ○ I have no difficulty walking up or down 12 stairs
    - ○ I have some difficulty walking up or down 12 stairs
    - ○ I have a lot of difficulty walking up or down 12 stairs
    - ○ I am not able to walk up or down 12 stairs
    - ○ Don't know/Don't remember

29. Continuing to think about the ***past week***, how much ***difficulty*** did you have rising from sitting?
    - ○ Extreme
    - ○ Severe
    - ○ Moderate
    - ○ Mild
    - ○ None
    - ○ Don't know/Don't remember

30. Continuing to think about the ***past week***, how much ***difficulty*** did you have standing?
    - ○ Extreme
    - ○ Severe
    - ○ Moderate
    - ○ Mild
    - ○ None
    - ○ Don't know/Don't remember

31. Continuing to think about the ***past week***, how much ***difficulty*** did you have getting on/off toilet?
    - ○ Extreme
    - ○ Severe
    - ○ Moderate
    - ○ Mild
    - ○ None
    - ○ Don't know/Don't remember

Section 5. About You

32. What is the highest grade or level of school that you completed?
    - ○ Some high school, but did not graduate
    - ○ High school graduate or GED
    - ○ Some college or 2-year degree
    - ○ 4-year college degree
    - ○ More than 4-year college degree
    - ○ I prefer not to answer

33. What was your total household income before taxes during the past 12 months?
    - ○ Less than $12,500
    - ○ $12,500-$19,999
    - ○ $20,000-$29,999
    - ○ $30,000-$49,999
    - ○ $50,000-$75,000
    - ○ Greater than $75,000
    - ○ I prefer not to answer

34. Are you of Hispanic, Latino, or Spanish origin?
    - ○ No, not of Hispanic, Latino, or Spanish origin
    - ○ Yes, of Hispanic, Latino, or Spanish origin
    - ○ I prefer not to answer

35. What is your race? **Choose all that apply**.
    - ○ White
    - ○ Black or African American
    - ○ American Indian or Alaska Native
    - ○ Asian
    - ○ Native Hawaiian or Other Pacific Islander
    - ○ I prefer not to answer

--- End of survey ----

2. Analytic Methodology

This section describes the general analytic approach for the survey of LEJR patients.

a. Survey Domains and Measures

We analyzed 20 survey measures, organized in five domains (functional status and pain, caregiver help, care transitions, care management, and overall recovery). The full list of measures is available in **Exhibit B-13**.

Exhibit B-13: Patient Survey Domains and Measures

Domain	Survey measures[a]	Description of survey measures
1. Functional status and pain[b]	Change in mobility	Ability to walk by yourself without resting
		Difficulty walking up or down 12 stairs
		Difficulty rising from sitting
		Difficulty standing
		Use of a mobility aid
	Change in toileting	Difficulty getting on or off the toilet
	Change in pain	Frequency that pain interferes with normal activities
	Change in medication	Medication use for pain in the joint you had replaced
2. Overall recovery	Satisfaction with overall recovery	Satisfaction with overall recovery since leaving the hospital
3. Care management	Composite measure of satisfaction with care management	Health care providers listened to preferences
		Satisfaction with discharge destination
		Satisfaction with care coordination
		Satisfaction with treatment instructions
4. Care transition	Discharged from the hospital at the right time	Discharged from the hospital at the right time
	Received the right amount of post-discharge care	Received the right amount of post-discharge care
	Had all the medical equipment needed at home	Had all the medical equipment needed at home
5. Caregiver help	Received any caregiver help	Received any caregiver help
	Composite measure of caregiver help	Help needed putting on or taking off clothes
		Help needed bathing
		Help needed using the toilet

Notes: LEJR = lower extremity joint replacement.

[a] Items regarding pain and medication referred directly to the joint that received surgery. All other items referred directly to the anchor hospitalization.

[b] For the eight functional status and pain measures, we modeled the change in functional status, where change was the difference between recalled status the week prior to the LEJR surgery, and reported status at the time the survey was completed.

Survey respondents were asked to recall their functional status and pain the week prior to their hospitalization, and to report their functional status and pain at the time of the survey, across eight related measures of function and pain. Each measure consisted of a Likert scale with three, four, or five levels. For each of the eight measures, we calculated the change in functional status or pain as the difference between a beneficiary's level of function or pain at the time of the survey and their recalled level of function or pain. We converted differences in levels of the Likert scale to percentage terms by dividing them by the average recalled level among CJR respondents. That is, the percentage difference was the difference between CJR and control respondents in relation to CJR respondents' recalled level of function or pain prior to their hospitalization.

In the caregiver help domain, measures of activities of daily living consisted of a Likert scale with three levels. Measures of satisfaction with care management or recovery consisted of a Likert scale with five levels. Results in these domains were normalized so that the lowest response category

(that is, "very dissatisfied" or "complete help needed") yielded a score of 0 and the highest response category (that is, "very satisfied" or "no help needed") yielded a score of 100.

The survey included three measures of care transition. The first measure, timing of discharge, had three response options (discharged too early, at the right time, or too late). The second measure, level of post-acute care received, included three response options (level of care during 2 weeks after surgery was more than respondent needed, about right, or not enough). The third measure, asking whether the respondent had all the medical equipment needed when sent home, had two response options (yes or no).

For all measures, we excluded respondents who did not respond to the specific item(s) used to construct a given measure. That is, we did not impute missing survey responses.

Composite Measure

We created composite measures for two domains: (1) "reliance on caregiver help (caregiver help)," conditional on having any caregiver help, summarized responses to three questions, and (2) "satisfaction with care management," which summarized responses to four questions. To create the composite measure, we translated response items into numeric scores and set them so that a score of 0 represented "the most negative amount of the construct being measured" (for example, most amount of caregiver help). Response categories were added so that the composite measure for a given domain was the sum of scores for its individual questions. For example, the "caregiver help" measure summarized three survey questions that each had three possible answers (0 – "complete help needed," 1 – "some help needed," or 2 – "no help needed"). The composite measure of "caregiver help" therefore ranged from 0, meaning maximum help needed, to 6, no help needed for any of the three tasks. Consistent with the Consumer Assessment of Health Providers and Systems® scoring, we rescaled the composite items so that scores ranged from 0 to 100, where zero indicated the least favorable outcome of the construct being measured (that is, greatest reliance on caregiver help and least satisfaction with care management).

b. Results Estimation

For each patient survey measure, we estimated the difference between CJR and control respondents.

For our analysis, we used the nonlinear model listed below for beneficiaries i, hospitals k, and wave t using a general functional form:

$$Y_{i,k,t} = f(b_1 + b_2 \cdot CJR_i + X_{i,k,t}' \cdot B + u_{i,k,t}) \quad \textbf{(3)}$$

Where:

Coefficient b_2 captured the difference in outcomes between CJR and control episodes,

$X_{i,k,t}$ indicated risk factors controlled for in our model, and f is a nonlinear function.

Functional forms for each measure are listed in **Exhibit B-14**.

Exhibit B-14: Functional Form for Regression Analyses

Domain	Survey measures	Functional form
1. Functional status and pain	All measures	Ordered logit
2. Overall recovery	Satisfaction with overall recovery	Logit
3. Care management	Composite measure of satisfaction with care management	Two-part model with logit first stage (any satisfaction) and negative binomial second stage (amount of satisfaction conditional on any)
	Individual satisfaction measures	Multinomial logit
4. Care transition	All measures	Logit
5. Caregiver help	Composite measure of caregiver help	Ordered logit
	Individual caregiver help measures	Multinomial logit

Risk Adjustment to Control for Differences in Patient Demographics and Clinical Risk Factors

All survey analyses controlled for potential differences in characteristics of the procedure, patient demographics and Medicare enrollment status, prior health conditions, and survey dimensions (first four domains in **Exhibit B-15**). We selected the patient-level characteristics as covariates for all survey analyses, based on the factors most strongly correlated with patient experience on the prior BPCI initiative patient survey and conceptual considerations (that is, factors predicted to be important based on theory).

Exhibit B-15: Risk Adjustment to Control for Differences in Patient Demographics and Clinical Risk Factors

Domain	Variables
1. Characteristics of the procedure	Fracture Knee replacement procedure MS-DRG (IP episodes) IP versus OP procedure
2. Patient demographics and Medicare enrollment status	Age Sex Dual Medicare or Medicaid eligibility Originally qualified for Medicare due to disability Assignment to ACO Self-reported race or ethnicity[a] Self-reported education[a] Self-reported prehospital functional status[a]
3. Prior health conditions	HCC score Stay in SNF or nursing home in 6 months prior to admission
4. Survey dimensions	Proxy status (patient had help from someone else in responding to the survey)

Domain	Variables
5. Optional patient, hospital, and MSA-level covariates[b]	Self-reported income Hospital size (staffed beds) Hospital academic affiliation Hospital ownership type Hospital prior BPCI experience (LEJR) Hospital prior BPCI experience (non-LEJR) PGP prior BPCI experience (LEJR) Hospital prior BPCI-A experience (non-LEJR) LEJR market competitiveness in MSA Medicare Advantage penetration in MSA (%)

Source: Risk adjustment variables were contructed from Medicare fee-for-service claims and beneficiary enrollment data, 2022 provider of service, Fiscal Year 2016 CMS Annual IPPS , CMS MDM, 2015-2016 AHRF, and BPCI initiative participant list.

Notes: ACO = Accountable Care Organization; BPCI = Bundled Payment for Care Improvement Initiative; BPCI-A = Bundled Payments for Care Improvement Advanced; HCC = hierarchical condition category; IP = inpatient; LEJR = lower extremity joint replacement; MSA = metropolitan statistical area; MS-DRG = Medicare Severity-Diagnosis Related Group; OP = outpatient; PGP = physician group practice; SNF = skilled nursing facility.

[a] For risk adjustment measures that are self-reported (that is, prehospital functional status; race or ethnicity; education), we coded all missing responses as 0 and included an additional binary variable indicating "missing item" (for example, missing race or ethnicity).

[b] While the first four domains acted as fixed covariates for our models, each measure's final risk-adjusted model included some unique combination of these optional variables, as well as squared and interaction terms.

In addition to the fixed variables, which we controlled for in all regressions, we ran a stepwise selection procedure on each outcome to test for additional control variables. Potential new variables included squared and interaction terms among the fixed variables, as well as the optional patient, hospital, and MSA-level variables listed in the fifth domain in **Exhibit B-15**.

Weighting

We employed entropy balancing to address potential differences in key patient characteristics across the CJR and control patients, and to mitigate potential differences between our sample of respondents and the populations from which they were drawn. The entropy balance weights minimized differences between the CJR and control patients on key attributes (see the first four domains in **Exhibit B-15**) and minimized differences in observable patient characteristics between CJR or control respondents relative to the full CJR population. For the analysis focused on respondents with hip fractures, we weighted the sample of CJR and control respondents to reflect the CJR population of beneficiaries with hip fractures. We applied weights to the full analytic sample and did not account for item nonresponse in calculating weights.[112]

Comparing Recalled Functional Status in the Week Prior to Hospitalization Between CJR and Control Respondents

We analyzed changes in self-reported functional status, and whether this differed for CJR and control respondents. We defined change as the difference between recalled status the week prior to the LEJR surgery and reported status at the time the survey was completed. Although we controlled for recalled prehospital functional status, our results may still be biased if CJR and

[112] Technically, the analytic sample varies slightly by question given item nonresponse, but we did not develop unique weights for each measure.

control respondents had substantially different functional status prior to surgery. For each of the eight prehospital functional status measures, we calculated the standardized difference in the unweighted mean between CJR and control respondents. Standardized differences for prehospital functional status between CJR and control respondents were nearly all close to or below 0.10 among elective episodes, and across all LEJRs, which is a conservative threshold for identifying potentially problematic differences between two groups (**Exhibit E-1, Exhibit E-2**).[113]

Comparing Claims-Based Patient Characteristics Between CJR and Control Respondents

Differences in functional status and pain between CJR and control respondents were risk-adjusted for a number of measures, including a fixed set of claims-based patient and episode characteristics (**Exhibit E-3, Exhibit E-4**). For each of these measures, we calculated the standardized difference in the unweighted mean between CJR and control respondents. Standardized differences were below 0.10 for all variables across all groups, with a few exceptions. CJR respondents with hip fractures were less likely than control respondents with hip fractures to have a major complication or comorbidity and were also less likely to be assigned to an ACO.

c. Sensitivity Analysis

Unlike CJR hospitals, CMS allowed control hospitals to enroll in BPCI Advanced for LEJR, a voluntary episode-based payment model, which could potentially introduce bias. If such bias existed, we would expect results to change if we dropped control episodes attributed to a BPCI Advanced LEJR. Hospitals participating in the BPCI Advanced LEJR clinical episode initiated 5.5% of all control episodes, and 14.0% of control episodes were attributed to physician group practices participating in the BPCI Advanced LEJR clinical episode. As a sensitivity analysis, we excluded all of these episodes, which did not change our results (see **Appendix E: Patient Survey Results**, **Exhibits E-10 to E-4**). This suggests that BPCI Advanced involvement within control hospitals did not bias our CJR patient survey results.

d. Limitations

The analyses had potential limitations related to the sample, timing of the survey, potential for recall bias, and differential characteristics of CJR and control respondents. More than one in three patients with an OP episode and nearly one in two patients with elective IP episodes did not respond to the survey, while roughly two in three hip fracture patients did not respond (**Exhibit B-12)**. Although we applied nonresponse weights to account for observable patient characteristics, to the extent non-respondents differed from respondents on unobservable factors correlated with our outcomes of interest, our results may not generalize to all patients in CJR. Since most survey measures focus on past events, incorrect recall may lead to mismeasurement of outcomes. This type of measurement error would not change the results, on average, because the same recall issue would have applied to both intervention and control groups, but it would reduce the precision of the estimates (that is, greater confidence intervals). We also note that there may be ceiling effects in some measures introduced by the fact that the majority of respondents indicated the highest levels of satisfaction with care management and recovery. Limited variation in these

[113] Austin, P. C. (2011). An introduction to propensity score methods for reducing the effects of confounding in observational studies. *Multivariate Behavioral Research, 46*(3), 399–424.

outcomes makes it more difficult to identify potentially meaningful differences in outcomes between the CJR and control groups.

G. Impact of the CJR Model on Underserved Populations

The CJR Model, which was implemented in 2016, did not explicitly include principles of health equity into the model design.[114] To study the impact of the CJR Model on underserved populations, we (1) measured differences, or "gaps," in health outcomes experienced by people disadvantaged or underserved at baseline and (2) examined whether CJR narrows or widens these gaps. While standard, this approach benchmarks the outcomes for an underserved population to the outcomes for a reference population. In race and ethnicity analyses, researchers typically use beneficiaries who are Non-Hispanic White as the reference population. Underserved populations may have systematically different health conditions and health care needs than reference populations. Consequently, the optimal level of a health outcome may differ for the underserved population and the reference population. For example, beneficiaries who are dually eligible for Medicare and Medicaid might have a higher level of appropriate emergency department visits than beneficiaries who are not dually eligible. It can be difficult to attach normative inferences to the gaps, or the changes in gaps, unless we know the "right" level of an outcome for a population.

For health equity analyses, we focused on the following four underserved populations:[115]

1. Patients who are dually eligible for Medicaid and Medicare
2. Patients who are Black or African American
3. Patients who are both dually eligible and Black or African American
4. Patients who are Hispanic

To study the impact of CJR, we conducted the following analyses:

1. Studied the proportion of the CJR (and control) populations that is underserved
2. Examined the impact of the model on the rate of LEJRs
3. Examined the direct impact of the model on payment per episode of care, utilization, quality of care, patient functional status and experience with care

1. Proportion of Underserved Populations in the CJR Model

We examined the proportion of the CJR population comprised of four different underserved population groups defined in **Exhibit B-16**. We defined these groups based on Medicare FFS beneficiary residence or receipt of LEJR in PY 6. Specifically, for each identified population

[114] Additional information about the CMS health equity framework is available at https://www.cms.gov/files/document/cms-framework-health-equity-2022.pdf.

[115] The analyses of the changes in race and ethnicity rely on the Research Triangle Institute (RTI) race codes from the Master Beneficiary Summary File. The RTI race code is created based on beneficiaries' self-reporting to Medicare and the Social Security Administration and RTI's race imputation algorithm utilizing beneficiary names and geographic locations. Additional information about the RTI race code can be found at https://resdac.org/cms-data/variables/research-triangle-institute-rti-race-code.

group, we calculated the proportion who are (1) Black or African American, (2) Hispanic, (3) dually eligible, and (4) both Black or African American and dually eligible (see **Exhibit B-17**).

Exhibit B-16 Proportion of Underserved Beneficiaries and Episodes

Population definition criteria	Population group definitions by CJR and control		Proportion calculation
	CJR	Control	
Beneficiary residing at CJR/control MSA in a year[a] (01/2022 – 12/2022)	• ***CJR MSA beneficiaries:*** Medicare FFS beneficiaries *residing* in a CJR MSA. • ***CJR MSA beneficiaries w/ LEJR:*** CJR MSA beneficiaries who had at least one LEJR. • ***CJR MSA beneficiaries w/ elective LEJR:*** CJR MSA beneficiaries who had at least one elective LEJR.	• ***Control MSA beneficiaries:*** Medicare FFS beneficiaries residing in a control MSA. • ***Control MSA beneficiaries w/ LEJR:*** Control MSA beneficiaries who had at least one LEJR. • ***Control MSA beneficiaries w/ elective LEJR:*** Control MSA beneficiaries who had at least one elective LEJR.	Denominator = Number of FFS beneficiaries for the population. Numerator = Among denominator, number of FFS beneficiaries identified in each underserved population based on RTI race code or dual eligibility indicator in Medicare enrollment data.
LEJR Episode[b] (Episode ending between 10/2021 and 12/2022)	• ***CJR Elective LEJR episodes:*** Elective LEJR episodes at a participating CJR hospital.	• ***Control Elective LEJR episodes:*** Elective LEJR episodes at a control hospital.	Denominator = Number of LEJR episodes. Numerator = Number of episodes related to patients identified in each underserved population based on RTI race code or dual eligibility indicator in Medicare enrollment data.

Source: CJR evaluation team analysis of Medicare claims and enrollment data for Medicare FFS beneficiaries enrolled in Medicare during 2022.

Notes: CJR = Comprehensive Care for Joint Replacement; FFS = fee-for-service; LEJR = lower extremity joint replacement; MSA = metropolitan statistical area; PY = performance year.

[a] Data source is a beneficiary/year-level analytic dataset capturing the list of all Medicare FFS beneficiaries, excluding beneficiaries eligible for Medicare based on end-stage renal disease status, living in mandatory CJR MSAs or their corresponding control MSAs in each calendar year between 2012 and 2022. The data include information on beneficiary sociodemographic characteristics as well as any LEJRs the beneficiary had in a year and were constructed from Medicare claims and enrollment data. For these data, the variables related to LEJR were constructed based on all LEJRs, including those that were performed at an ambulatory surgical center. For this analysis, we only used data for the Calendar Year 2022, which aligns with most of PY 6.[116]

[b] Data source is an episode-level analytic dataset capturing information on all episodes of care related to LEJRs performed at a participating CJR hospital or their corresponding control hospital between baseline and PY 6 (2012–2022). Note that some hospitals in the CJR MSAs are not participating in the model due to low-volume or rural designation starting from PY 6. Only episodes related to hospitals participating in the CJR Model in PY 6 have been included. For this analysis, we only used data related to PY 6. Some LEJR episodes at CJR participating hospitals can be for Medicare FFS beneficiaries who are not residing in a CJR MSA, whereas some beneficiaries who live in a CJR MSA could have had their LEJR episode at a nonparticipating CJR hospital. Thus, the counts generated from beneficiary/year data and LEJR episode data are slightly different.

[116] While PY 6 includes episodes ending between October 2021 and December 2022, the beneficiary-level analytic dataset captures information in a calendar year. Thus, there are episodes that started in the last 2 quarters of 2021 that are PY 6 episodes and not included in the beneficiary-level dataset for 2022, and there are LEJRs present in the beneficiary-level dataset that occurred in the fourth quarter of 2022 that are not included in PY 6.

Some LEJR episodes at CJR participating hospitals can be for Medicare FFS beneficiaries who are not residing at a CJR MSA while some beneficiaries who live in a CJR MSA could have had their LEJR episode at a non-participating CJR hospital. Hence the counts generated from beneficiary/year data and LEJR episode data are slightly different.

Exhibit B-17 Sample Sizes for Analysis of Proportion of Underserved Beneficiaries and Episodes

Sample sizes	Medicare FFS beneficiaries in 2022[a]		Medicare FFS beneficiaries with an LEJR in 2022[a]		Medicare FFS beneficiaries with an elective LEJR in 2022[a]		PY 6 episodes[b]	
	CJR MSAs	Control MSAs	CJR MSAs	Control MSAs	CJR MSAs	Control MSAs	CJR hospitals	Control hospitals
Total population	4,504,861	4,348,736	68,652	75,039	60,231	66,474	47,070	51,857
Black or African American	360,825	391,452	3,200	3,930	2,902	3,608	1,946	2,668
Hispanic	378,628	215,328	3,264	1,948	2,877	1,747	2,315	1,341
Dually eligible	852,884	586,533	6,003	4,537	4,307	3,436	3,108	2,616
Black or African American and dually eligible	131,534	132,208	709	863	579	721	396	484

Source: CJR evaluation team analysis of Medicare claims and enrollment data for Medicare FFS beneficiaries enrolled in Medicare during 2022.

Notes: CJR = Comprehensive Care for Joint Replacement; FFS = fee-for-service; LEJR = lower extremity joint replacement; MSA = metropolitan statistical area; PY = performance year.

[a] Data source is a beneficiary/year-level analytic dataset capturing the list of all Medicare FFS beneficiaries, excluding beneficiaries eligible for Medicare based on end-stage renal disease status, living in mandatory CJR MSAs or their corresponding control MSAs in each calendar year between 2012 and 2022. The data include information on beneficiary sociodemographic characteristics as well as any LEJRs the beneficiary had in a year and were constructed from Medicare claims and enrollment data. For these data, the variables related to LEJR were constructed based on all LEJRs, including those that were performed at an ambulatory surgical center. For this analysis, we only used data for the Calendar Year 2022, which aligns with most of PY 6.

[b] Data source is an episode-level analytic dataset capturing information on all episodes of care related to LEJRs performed at a participating CJR hospital or their corresponding control hospital between baseline and PY 6 (2012–2022). Note that some hospitals in the CJR MSAs are not participating in the model due to low-volume or rural designation starting from PY 6. Only episodes related to hospitals participating in the CJR Model in PY 6 have been included. For this analysis, we only used data related to PY 6.

2. *Analysis of the Impact of CJR Model on Claims-Based Outcomes for Underserved Populations*

Our analysis of the differential impact of the CJR Model on underserved subpopulations was based on the DiD methodology. We studied the differential impact of the CJR Model by estimating the impact of the CJR Model on an underserved subpopulation and a reference subpopulation, and then we estimated the difference between the two CJR Model impacts to determine whether the CJR Model affected the underserved subpopulation differently from the reference subpopulation. The estimation of both differential impacts took place in a single regression, subject to the constraint that the coefficients on risk-adjustment variables were the same for both subpopulations.

As noted above, we studied four underserved subpopulations: (1) patients who are Black or African American, (2) patients who are eligible for both Medicare and Medicaid (dually eligible), (3) patients who are Black or African American *and* dually eligible, and (4) patients who are

Hispanic. The reference subpopulations were, respectively, (1) patients who are Non-Hispanic White, (2) patients who were not dually eligible, (3) patients who both are Non-Hispanic White and not dually eligible, and (4) patients who are Non-Hispanic White. The estimated differential impact represented how the difference in the risk-adjusted average outcome between the underserved subpopulation (for example, patients who are Black or African American) and the reference subpopulation (for example, patients who are Non-Hispanic White) changed between the baseline and intervention periods due to the CJR Model. In other words, it captured the difference between the effect of the CJR Model on the underserved subpopulation and the effect of the CJR Model on the reference subpopulation, also known as a triple difference (difference-in-difference-in-differences, or DDD) estimate.

In general, the statistical model to estimate the DDD was:

$$Y_{i,k,t} = b_0 + b_1 \cdot t + b_2 \cdot CJR_{i,k} + b_3 \cdot G_{i,t} + b_4 \cdot G_{i,t} \cdot t + b_5 \cdot CJR_{i,k} \cdot t + b_6 \cdot CJR_{i,k} \cdot G_{i,t} \cdot t + X_{i,k,t}{}' \cdot B + u_{i,k,t} \quad \textbf{(4)}$$

- Yi,k,t was the outcome for the ith episode with an LEJR at hospital k in period t (t = 1 during the CJR PY 6 intervention quarters and 0 otherwise).
- $CJR_{i,k}$ was an indicator that takes the value of 1 if the i^{th} episode was initiated by a CJR participant hospital *k* and takes the value of 0 otherwise.
- $G_{i,t}$ was an indicator that takes the value of 1 if the patient for the i^{th} episode belongs to the underserved subgroup and takes the value of 0 if the patient belonged to the reference subgroup.
- $X_{i,k,t}$ was hospital, geographic, and patient characteristics in period *t.*
- The value of coefficient b_1 captured aggregate factors that could cause changes in outcome *Y* in the intervention period relative to the baseline period that are common across CJR and control group episodes.
- Coefficient b_2 captured the relative differences in outcomes between CJR and control group episodes.
- Coefficient b_3 captured the relative differences in outcomes between patients in the underserved subgroup and patients in the reference group.
- Coefficient b_4 captured aggregated factors that could cause changes in outcome *Y* for underserved subgroup $G_{i.t}$ in the intervention period relative to the baseline period that are common across CJR and control group episodes.
- Coefficient b_5 determined the differential in outcome *Y* experienced by patients receiving services from CJR hospitals during the CJR intervention period relative to control group episodes in the intervention period and represented the DiD estimator.
- Coefficient b_6 determined the differential in outcome *Y* experienced by patients in underserved subgroup $G_{i.t}$ receiving services from CJR hospitals during the CJR intervention period relative to control group episodes in the intervention period and represented the DDD estimator.
- The vector of coefficients *B* measured the differential effects of risk factors (*X*) on the outcome variable.

We explored differential impacts of the CJR Model on eight outcomes: total payments, HHA visits, SNF length-of-stay, and IRF length-of-stay used OLS regressions. For mortality, emergency department use, and unplanned readmission rates, we used logistic regression. Finally, for first post-acute care discharge destination, we used a multinomial logistic regression.

3. Analysis of the Impact of CJR Model on LEJR Volume for Underserved Populations

Prior research has reported widening differences in access to LEJR for certain underserved populations under the CJR Model.[117,118] We evaluated the impact of the CJR Model on the volume of LEJR discharges for four underserved populations.

a. Sample and Time Periods

Our sample included all Medicare FFS beneficiaries who resided in a mandatory CJR or control MSA in either the baseline or PY 6. Because PY 6 began in October 2021, the middle of a calendar year, our analyses using the beneficiary-year sample excluded 2021 data and focused on 2022 instead. We made a similar change to the baseline period for this analysis. Rather than running from January 2012 to March 2015, ending in the middle of a calendar year, we extended the baseline end date to December 2015.

Where appropriate, the sample exclusion rules followed those for our main analyses (for instance, excluding beneficiaries who were eligible for Medicare based on ESRD status). In addition, the beneficiary-year sample included a beneficiary-year observation only if the beneficiary was also eligible for inclusion in the prior year. The resultant analytic sample included 43,168,817 beneficiary-years.

b. LEJRs per 100,000 Fee-for-Service Population

We present results in terms of the rate of LEJRs per 100,000 FFS beneficiaries per year. We calculated this rate by multiplying the estimated probabilities from the DiD or DDD analysis by 100,000.

c. Statistical Model (overall difference-in-differences)

The outcome of interest was whether a patient received at least one elective LEJR in a given year. We used a logistic regression model, which incorporated controls for beneficiary characteristics, HCC flags, prior care use, and state fixed effects.

$$logit(Y_{it}) = b_0 + b_1\boldsymbol{C}_{it} + b_2\boldsymbol{HCC}_{it} + b_3\boldsymbol{PC}_{it} + b_4 CJR_{it} + b_5 After_t + b_6(CJR_{it} \times After_t) + b_7\boldsymbol{S}_{it} + \varepsilon_{it}$$

- Y_{it} was an indicator variable that takes on the value of 1 if beneficiary *i* received at least one elective LEJR in year *t*.

[117] Thirukumaran, C., Kim, Y., Cai, X., Ricciardi, B., Li, Y., Fiscella, K., Mesfin, A., & Glance, L. (2021). Association of the Comprehensive Care for Joint Replacement Model with disparities in the use of total hip and total knee replacement. *JAMA Network Open, 4*(5), e2111858.

[118] Kim, H., Meath, T., Quiñones, A., McConnell, J., & Ibrahim, S. (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, 4*(3), e211772.

- $\boldsymbol{C}_{it}$ was a vector of beneficiary characteristics, including indicator variables for race, dual eligibility for Medicare and Medicaid, age buckets, sex, and certain chronic complications, such as obesity and diabetes.
- $\boldsymbol{HCC}_{it}$ was a vector of indicator variables for a subset of HCC flags.
- $\boldsymbol{PC}_{it}$ was a vector of variables for prior care use—specifically, continuous variables reflecting the number of days receiving health care services in the 6 months prior to the LEJR.
- $\boldsymbol{S}_{it}$ was a vector of binary variables indicating the state in which the beneficiary resided.
- CJR_{it} was an indicator for whether a patient resided in a CJR MSA at time *t*.
- $After_t$ was an indicator for whether the episode occurred during the baseline (=0) or during the last 4 quarters of PY 6, in 2022 (=1).

The impact of the CJR Model was captured by the coefficient b_6, which was identified by comparing the beneficiaries who resided in CJR MSAs during PY 6 ($After_t = 1$) to beneficiaries who resided in CJR MSAs during the baseline period ($After_t = 0$) and then comparing that difference to the same difference calculated on beneficiaries who resided in control MSAs. We then transformed the coefficient into an LEJR rate, the number of LEJRs performed per 100,000 FFS beneficiaries per year. Standard errors were clustered at the MSA level.

d. Statistical Model (health equity difference-in-difference-in-differences)

Conducting the DDD analysis was similar to performing a DiD on two subpopulations and then comparing the resulting estimates.

The model specification for the health equity DDD analyses was:

$$logit(Y_{it}) = b_0 + b_1\boldsymbol{X}_{it} + b_2G_{it} + b_3CJR_{it} + b_4After_t + b_5(G_{it} \times CJR_{it}) + b_6(G_{it} \times After_t) + b_7(CJR_{it} \times After_t) + b_8(G_{it} \times CJR_{it} \times After_t) + \varepsilon_{it}$$

Y_{it} was an indicator variable that takes on the value of 1 if beneficiary *i* received at least one elective LEJR in year *t*, and $\boldsymbol{X}_{it}$ was a vector of control covariates, containing all of the covariates mentioned in **Section C**, above. G_{it} was a binary variable that indicated whether beneficiary *i* in year *t* was a member of the underserved subpopulation or the reference subpopulation. The coefficient b_8 captured the difference between the estimated effect of the CJR Model on the underserved subpopulation and on the reference subpopulation. For instance, if G_{it} was 1 if a beneficiary was dually eligible, and 0 otherwise, then b_8 would capture the difference in the estimated effect of the CJR Model on the probability of receiving an LEJR between beneficiaries with dual eligibility and beneficiaries without dual eligibility.

e. Limitations

An important limitation to the beneficiary-year analyses was that the sample, by construction, does not capture LEJRs performed on beneficiaries who did *not* reside in a mandatory CJR or control MSA but who received LEJRs in mandatory CJR or control MSAs. A non-negligible portion of LEJRs performed in mandatory CJR or mandatory control MSAs involved beneficiaries traveling from other locations, and these LEJRs were not captured in the beneficiary-year analysis. While a

smaller population, beneficiaries who reside in the CJR or control MSAs who received LEJRs in hospitals that were not in these areas are not included in this analysis.

H. Impact of the CJR Model on Functional Status and Care Experiences for Underserved Populations

Our main patient survey analysis measured the impact of CJR on the average patient. However, we are particularly concerned about changes in experiences of patients who rate care as poor, and changes in patient experience that move patients to worse, or better experiences. This is especially salient for underserved patients who often differ from the average patient in health and health care. In this analysis, we looked at the differences between CJR and control patients across the distribution of responses of survey measures. The analysis allowed us to better study heterogeneous results for underserved populations, and specifically detect improvement or degradation in care for patients who may not reflect the "average" beneficiary. (See **Chapter III: Impact of the Model** for additional information about the patient survey.)

1. Survey Domains and Measures

We analyzed the individual survey measures of functional status and pain, caregiver help, care management, and overall recovery described in **Exhibit B-13** in **Appendix B: Data and Methods**. We did not use survey composite measures for this analysis.

2. Sample and Time Periods

The sample was restricted to Wave 8 of the patient survey, which took place during PY 6. We studied three underserved populations: (1) patients who are dually eligible, compared with patients who are *not* dually eligible for Medicaid; (2) patients who are Black or African American, compared with patients who are Non-Hispanic White; and (3) patients who are Hispanic, compared with patients who are Non-Hispanic White.

Our dataset contained 9,676 survey respondents; among these survey respondents, 402 are LEJR patients who were dually eligible for Medicaid, 370 are LEJR patients who are Black or African American, and 307 are LEJR patients who are Hispanic. These numbers were lower than might be expected because a high rate of respondents chose not to provide their self-identified race or ethnicity.

3. Statistical Model

The statistical model used in our health equity analysis of the patient survey data differed from that used in the main analysis of these data. The main analysis treated ordered responses and changes in ordered responses as cardinal numbers. This was a valid approach to estimating the *mean* effect of the CJR Model. However, subpopulation effects may have been subtle and heterogeneous. Consequently, for our health equity analysis, we estimated effects for each possible response. Specifically, we treated ordered responses as ordinal data and used the method of ordinal logistic regression. The regression specification was:

$$Pr(outcome_i = j) = Pr\left(\kappa_{j-1} < b_1 \boldsymbol{X}_{it} + b_2 CJR_i + b_3 \boldsymbol{G_i} + b_4 (\boldsymbol{G_i} \times CJR_i) + \varepsilon_i \leq \kappa_j\right)$$

The probability that patient *i* chooses outcome *j* was the probability that a linear function plus a random error was within a range of cut points for the outcome. $\boldsymbol{G}_i$ was a vector of binary variables

that indicated the subpopulation identities of patient *i*. The random error was assumed to be logistically distributed. The coefficients of the linear function and the values of the cut points were estimated by maximum likelihood. The regression was weighted using the entropy-balancing weights described in **Section F.2.b** above and used robust standard errors for inference.

4. Qualitative Analysis

a. ACO Interviews

Between September and November 2023, we conducted telephone interviews with 32 CJR participant hospitals. The aim of the interviews was to understand CJR hospitals' experience with ACOs, especially Medicare ACOs.

Interviewees

We interviewed representatives from CJR participant hospitals including CJR service line coordinators, ACO representatives, population health department directors, and data analysts.

Protocol

We developed and implemented a 45-minute semi-structured interview guide that was tailored to answer the following key questions:

- What does it mean for a CJR hospital to be a part of a Medicare ACO?
- How are CJR hospitals influenced by their participation in a Medicare ACO?
- Are actions taken to respond to Medicare ACOs similar or different than those taken to respond to the CJR Model?
- What are the benefits and challenges of participating in a Medicare ACO and the CJR Model concurrently?
- What are the experiences with ACOs for CJR hospitals that do not participate in a Medicare ACO?

2021 AHA Survey

We determined the ACO participation status of hospitals in the CJR Model using the 2021 AHA Survey. Hospitals responded to the AHA Survey indicating whether they were leading a Medicare ACO, were leading another type of ACO, were participating but not leading in any type of ACO, or had never participated in an ACO. Hospitals that responded that they were leading a Medicare ACO were included in our sample as "Medicare ACO participants." Hospitals that responded they had never participated in an ACO, were leading another type of ACO, or were participating but not leading any type of ACO[119] were categorized as non-Medicare ACO hospitals in our sample.

[119] The language of the AHA Survey question on participation in any type of ACO made it unclear whether a hospital responding "Yes" was participating in a Medicare ACO or another type, so we considered these respondents as part of the non-Medicare ACO group for our sampling purposes.

Sample Selection

We selected a sample of 100 CJR hospitals. Fifty of the hospitals were leading a Medicare ACO, 25 hospitals were participating in but not leading an ACO or leading a non-Medicare ACO, and 25 hospitals were not participating in an ACO. We used the following criteria to select the sample.

- Inclusion criteria:
  - Responded (yes or no) to ACO participation question in the AHA Survey
  - Hospital had at least 20 LEJR episodes in PY 6
- Exclusion criteria:
  - No response to the ACO participation question in the AHA Survey data
  - Hospital had fewer than 20 LEJR episodes in PY 6
  - Hospital was part of health system with three CJR hospitals already included in the sample (used to ensure health systems with many CJR participant hospitals were not overrepresented in our study)

Recruitment

We conducted outreach to hospitals in our sample by sending an initial email invitation, asking whether they would participate in this round of interviews. We attached a frequently asked questions document and informed consent information to the email. Once the hospital responded, we followed up with a confirmation email and requested that the hospital point of contact complete the preinterview survey about the hospital's past ACO experience to inform and tailor the interview. If a hospital's reported ACO experience on the preinterview survey or during the interview differed from the AHA Survey, the hospital interview was categorized in the participation group that the hospital reported.

We conducted outreach in three waves to monitor our response rate and adjust outreach strategy if needed. Across the first two waves of outreach, we invited 65 hospitals to our interviews. Hospital response rate was higher than anticipated during the first two waves of randomly selected hospitals from our sample, and we achieved saturation for non-ACO participant hospitals. Therefore, the last wave was a smaller (eight hospitals) and a purposively sampled selection of AHA-surveyed Medicare ACO participant hospitals to ensure we had adequate representation from that group of hospitals.

Data Collection

We interviewed representatives from 32 hospitals—21 hospitals that participated in a Medicare ACO at the time of the interview and 11 hospitals that did not participate in a Medicare ACO at the time of the interview. We took notes during telephone interviews, and if the interviewee agreed, we recorded the interview. We used the recordings to verify and enhance interview notes. Notes from telephone interviews were organized and entered into ATLAS.ti software (version 23; Scientific Software Development GmbH, Berlin, Germany) for coding and analysis.

Analysis

Notetakers and interviewers who participated in interviews completed a thematic analysis of telephone interview data. We developed analytic codebooks including primary and sub-codes based on the telephone interview protocols. Coders used ATLAS.ti to apply codes and sub-codes to comprehensive interview notes and ran queries to identify themes across interviews. All coders received systematic training, which included parallel coding and discussion of results with trainers until consistency was established. Throughout the analysis, we refined the codebooks (that is, codes were dropped, consolidated, added, or revised) to better capture patterns as they emerged.

b. Safety-Net Hospital Interviews

We conducted telephone interviews in November and December 2023 with six CJR participant hospitals that we identified as being SNHs.

Interviewees

We interviewed representatives from CJR participant hospitals, including hospital staff and leadership, orthopedic surgeons, and system-level representatives.

Protocol

We developed and implemented a 60-minute semi-structured interview guide that was tailored to answer the following key questions:

- How do SNHs care for LEJR patients with unmet nonmedical needs?
- How does the CJR Model impact SNHs?
- What do SNHs need to better care for LEJR patients with unmet nonmedical needs?

Sample Selection

We selected a sample of 17 SNHs, which were identified using three attributes:

1. Disproportionate Share Hospital percentage (DSH%)
2. Dual eligibility percentage (Dual%)
3. Area Deprivation Index

We excluded hospitals with fewer than 20 LEJR episodes and hospitals that were selected for the ACO interview sample (discussed above) from the SNH sample.

We categorized hospitals as being an SNH if they were in the 85th percentile or higher for DSH%. The sample was then selected based on the highest Dual% and DSH%, with consideration for geographic variability to ensure all the sampled hospitals were not from a single MSA. We identified 17 hospitals in total for the sample, with a goal of conducting interviews with six hospitals.

Recruitment

We conducted outreach to 11 hospitals in our sample by sending an initial email invitation asking whether they would participate in this round of interviews. We attached a frequently asked

questions document and informed consent information to the email. Once the hospital responded, we followed up with a confirmation email. Six of the 11 hospitals agreed to participate in our interviews, so we did not continue outreach to the remaining hospitals in our sample.

Data Collection

We interviewed representatives from six SNHs. We took notes during telephone interviews, and if the interviewee agreed, we recorded the interview. We used recordings to verify and enhance interview notes. Notes from telephone interviews were organized in notetaking templates. Notetakers and interviewers worked together to clean and finalize notes and prepare them for analysis.

Analysis

Notetakers and interviewers who participated in interviews completed a thematic analysis of telephone interview data. We developed analytic codebooks including primary and sub-codes based on the telephone interview protocols. Coders used ATLAS.ti to apply codes and sub-codes to comprehensive interview notes and ran queries to identify themes across interviews. All coders received systematic training, which included parallel coding and discussion of results with trainers until consistency was established. We refined the codebooks throughout the analysis (that is, codes were dropped, consolidated, added, or revised) to better capture patterns as they emerged.

Appendix C: CJR Population Patient Characteristics

A. Descriptive Tables

Exhibit C-1: Patient Characteristics for CJR and Control Populations in Both PY 6 and in the Baseline

Domain	Measure	CJR		Control	
		PY 6	Baseline	PY 6	Baseline
Demographics	Non-Hispanic White	85.4%	84.7%	88.5%	88.3%
	Black or African American	4.0%	5.9%	4.9%	7.1%
	Asian	2.1%	1.8%	1.3%	0.9%
	Hispanic	4.9%	6.3%	2.6%	2.8%
	Dually Eligible	7.8%	13.7%	5.7%	10.7%
	Female	63.6%	66.1%	63.4%	65.6%
Health conditions	Dementia	5.5%	7.6%	5.2%	7.1%
	Diabetes	27.2%	29.5%	24.9%	27.3%
	Congestive Heart Failure	14.0%	14.8%	13.3%	13.6%
	Hypertension	76.5%	75.2%	76.2%	75.4%
	Obesity	33.4%	15.3%	34.4%	16.4%
Prior care	Any Home Health	9.9%	13.2%	8.7%	12.2%
	Any Hospice	0.2%	0.2%	0.2%	0.2%
	Any Inpatient Prospective Payment System	7.5%	13.0%	7.2%	12.8%
	Any Inpatient Rehabilitation facility	0.7%	1.4%	0.9%	1.4%
	Any Skilled Nursing Facility	2.5%	5.1%	2.2%	4.4%

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 ending between October 1, 2021, and December 31, 2022 (PY 6 intervention).

Notes: Prior Care is defined as utilization six months prior to the anchor begin date.

CJR = Comprehensive Care for Joint Replacement Model; MSA = metropolitan statistical area; PY = performance year.

B. Underserved Population Patient Mix Tables

Exhibit C-2: Changes in Patient Mix for Patients Who Are Black or African American Relative to Patients Who Are Non-Hispanic White, Elective LEJRs, Baseline Compared With PY 6

Domain	Measure	Change from baseline to PY 6 for CJR vs. change from baseline to PY 6 for control								Difference in relative changes			
		Patients who are Black or African American				Patients who are Non-Hispanic White				Estimate of difference (pp)	p-value	90% LCI	90% UCI
		Estimate (pp)	p-value	90% LCI	90% UCI	Estimate (pp)	p-value	90% LCI	90% UCI				
Age	<65	2.8	0.28	-1.4	6.9	0.1	0.90	-0.9	1.1	2.7	0.23	-1.0	6.3
	65-74	-1.5	0.40	-4.4	1.4	1.0	0.45	-1.1	3.1	-2.4	0.23	-5.8	0.9
	75-84	-1.3	0.46	-4.1	1.6	-0.2	0.79	-1.7	1.2	-1.0	0.64	-4.6	2.6
	85+	-0.0	0.95	-0.9	0.9	-0.8	0.04	-1.5	-0.2	0.8	0.26	-0.4	1.9
Demographics	Female	-0.1	0.93	-2.4	2.2	0.2	0.70	-0.7	1.1	-0.3	0.80	-2.6	1.9
	Eligible for Medicaid	-1.1	0.69	-5.6	3.4	0.2	0.68	-0.7	1.2	-1.3	0.59	-5.4	2.7
	Disability, no ESRD	3.3	0.10	0.0	6.7	0.2	0.77	-1.0	1.4	3.1	0.06	0.4	5.8
Hierarchical Condition Categories (HCC)	Score a	0.0	0.64	-0.1	0.1	0.0	0.10	-0.1	0.0	0.0	0.76	-0.1	0.1
	Count b	-0.1	0.60	-0.3	0.2	-0.1	0.11	-0.2	0.0	0.0	0.76	-0.1	0.2
	Count: 0	2.5	0.16	-0.4	5.5	2.0	0.02	0.6	3.4	0.5	0.71	-1.8	2.8
	Count: 1	0.6	0.70	-2.1	3.3	0.9	0.13	-0.1	1.8	-0.3	0.86	-2.8	2.3
	Count: 2	-0.8	0.55	-3.0	1.4	-0.6	0.07	-1.1	-0.1	-0.3	0.85	-2.4	1.9
	Count: 3	1.6	0.35	-1.2	4.4	-1.0	0.02	-1.6	-0.3	2.6	0.10	-0.0	5.1
	Count: 4+	-3.9	0.10	-7.9	0.1	-1.3	0.24	-3.3	0.6	-2.6	0.23	-6.1	1.0
Health status	Obesity	-1.3	0.65	-6.0	3.4	0.4	0.88	-3.7	4.4	-1.7	0.42	-5.1	1.7
	Diabetes	-3.4	0.14	-7.1	0.4	0.2	0.77	-0.9	1.2	-3.5	0.08	-6.9	-0.2
	Hypertension	-0.4	0.80	-3.3	2.4	0.7	0.47	-0.9	2.4	-1.2	0.37	-3.4	1.0
	Dementia	0.2	0.78	-0.8	1.2	-0.0	0.94	-0.3	0.3	0.2	0.77	-0.8	1.2
	Congestive heart failure	1.7	0.28	-0.9	4.3	-0.5	0.40	-1.5	0.5	2.2	0.12	-0.1	4.6

Domain	Measure	Change from baseline to PY 6 for CJR vs. change from baseline to PY 6 for control								Difference in relative changes			
		Patients who are Black or African American				Patients who are Non-Hispanic White				Estimate of difference (pp)	p-value	90% LCI	90% UCI
		Estimate (pp)	p-value	90% LCI	90% UCI	Estimate (pp)	p-value	90% LCI	90% UCI				
Prior use	ACH stay	1.2	0.18	-0.3	2.7	0.4	0.38	-0.3	1.0	0.9	0.31	-0.5	2.3
	HH use	4.4	0.07	0.4	8.4	0.2	0.79	-0.8	1.2	4.2	0.05	0.7	7.8
	IRF stay	0.0	0.96	-0.6	0.6	-0.2	0.25	-0.5	0.1	0.2	0.51	-0.3	0.7
	SNF stay	-0.5	0.35	-1.4	0.4	-0.2	0.34	-0.6	0.2	-0.3	0.57	-1.1	0.5
	Any prior care	5.5	0.02	1.6	9.5	-0.0	0.96	-1.2	1.1	5.6	0.02	1.8	0.1

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 October 2021 that ended by December 2022 (intervention).

Notes: The estimates in this exhibit are the results of unadjusted DiD models on patient characteristics. The estimate of difference is the difference of the unadjusted DiD estimates. Estimates that are significant at the 1%, 5%, or 10% significance level are indicated by red, orange, or yellow shaded cells, respectively. Count characteristics refer to the count of individual HCCs. Any prior care included inpatient hospital, psychiatric hospital, emergency department, SNF, IRF, home health, long-term care hospital, and hospice during the 6 months prior to anchor hospitalization.

ACH = acute care hospital; DDD = triple difference-in-differences; DiD = difference-in-differences; ESRD = end-stage renal disease; HH = home health; IRF = inpatient rehabilitation facility; LCI = lower confidence interval; LEJR = lower extremity joint replacement; pp = percentage point; PY = performance year; SNF = skilled nursing facility; UCI = upper confidence interval.

[a] Estimates for HCC score are expressed as units of the score instead of percentage points.

[b] Estimates for the count of HCCs are expressed as numbers instead of percentage points.

Exhibit C-3: Changes in Patient Mix for Patients Who Were Dually Eligible Relative to Patients Who Were Not Dually Eligible, Elective LEJRs, Baseline Compared With PY 6

Domain	Measure	Change from baseline to PY 6 for CJR vs. change from baseline to PY 6 for control								Difference in relative changes			
		Patients who were dually eligible				Patients who are non-dually eligible				Estimate of difference (pp)	p-value	90% LCI	90% UCI
		Estimate (pp)	p-value	90% LCI	90% UCI	Estimate (pp)	p-value	90% LCI	90% UCI				
Age	<65	3.3	0.13	-0.2	6.8	0.0	0.97	-0.7	0.7	3.3	0.13	-0.3	6.8
	65-74	-1.6	0.40	-4.7	1.5	0.9	0.45	-1.0	2.8	-2.5	0.27	-6.2	1.2
	75-84	-2.2	0.17	-4.9	0.4	-0.2	0.83	-1.6	1.3	-2.0	0.32	-5.4	1.4
	85+	0.5	0.53	-0.8	1.9	-0.7	0.07	-1.3	-0.1	1.2	0.18	-0.3	2.7
Demographics	Female	3.1	0.02	1.0	5.3	-0.1	0.87	-1.0	0.8	3.2	0.05	0.6	5.9
	Black or African American	0.6	0.78	-3.1	4.4	-0.1	0.85	-0.7	0.5	0.7	0.74	-2.8	4.2
	Hispanic	-3.5	0.23	-8.4	1.4	0.1	0.69	-0.3	0.5	-3.6	0.21	-8.4	1.1
	Disability, no ESRD	2.0	0.44	-2.2	6.1	0.1	0.78	-0.7	1.0	1.8	0.47	-2.4	6.0
Hierarchical Condition Categories (HCC)	Score [a]	-0.0	0.87	-0.1	0.1	-0.0	0.09	-0.1	0.0	0.0	0.54	-0.1	0.1
	Count [b]	0.0	0.89	-0.2	0.2	-0.1	0.12	-0.2	0.0	0.1	0.23	0.0	0.3
	Count: 0	0.3	0.76	-1.5	2.2	2.0	0.02	0.6	3.5	-1.7	0.12	-3.5	0.1
	Count: 1	0.4	0.70	-1.4	2.2	0.9	0.14	-0.1	2.0	-0.5	0.64	-2.3	1.3
	Count: 2	-0.6	0.58	-2.4	1.2	-0.5	0.11	-1.0	0.0	-0.1	0.93	-2.1	1.8
	Count: 3	-0.9	0.48	-3.1	1.2	-1.0	0.02	-1.7	-0.3	0.1	0.96	-1.8	1.9
	Count: 4+	0.7	0.71	-2.6	4.1	-1.5	0.21	-3.5	0.5	2.3	0.27	-1.1	5.7
Health status	Obesity	-0.6	0.82	-5.1	3.9	0.1	0.97	-4.1	4.2	-0.7	0.68	-3.6	2.2
	Diabetes	0.0	1.00	-4.1	4.1	0.1	0.88	-1.0	1.2	-0.1	0.97	-4.0	3.8
	Hypertension	-0.2	0.93	-2.8	2.5	0.6	0.57	-1.1	2.3	-0.7	0.67	-3.6	2.1
	Dementia	-0.0	0.97	-1.3	1.3	0.1	0.46	-0.1	0.4	-0.1	0.85	-1.4	1.1
	Congestive heart failure	0.8	0.55	-1.4	2.9	-0.5	0.43	-1.5	0.5	1.2	0.28	-0.7	3.1

Domain	Measure	Change from baseline to PY 6 for CJR vs. change from baseline to PY 6 for control								Difference in relative changes			
		Patients who were dually eligible				Patients who are non-dually eligible				Estimate of difference (pp)	p-value	90% LCI	90% UCI
		Estimate (pp)	p-value	90% LCI	90% UCI	Estimate (pp)	p-value	90% LCI	90% UCI				
Prior use	ACH stay	1.1	0.32	-0.7	3.0	0.3	0.36	-0.3	0.9	0.8	0.46	-1.0	2.6
	HH use	4.0	0.21	-1.3	9.2	0.3	0.66	-0.7	1.3	3.7	0.22	-1.3	8.7
	IRF stay	-0.3	0.43	-0.9	0.3	-0.1	0.37	-0.4	0.1	-0.1	0.67	-0.6	0.4
	SNF stay	-0.4	0.52	-1.6	0.7	-0.2	0.56	-0.6	0.3	-0.3	0.73	-1.6	1.0
	Any prior care	5.0	0.13	-0.4	10.4	0.0	0.99	-1.0	1.0	5.0	0.11	-0.1	10.1

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 October 2021 that ended by December 2022 (intervention).

Notes: The estimates in this exhibit are the results of unadjusted DiD models on patient characteristics. The estimate of difference is the difference of the unadjusted DiD estimates. Estimates that are significant at the 1%, 5%, or 10% significance level are indicated by red, orange, or yellow shaded cells, respectively. Count characteristics refer to the count of individual HCCs. Any prior care included inpatient hospital, psychiatric hospital, emergency department, SNF, IRF, home health, long-term care hospital, and hospice during the 6 months prior to anchor hospitalization.

ACH = acute care hospital; DDD = triple difference-in-differences; DiD = difference-in-differences; ESRD = end-stage renal disease; HH = home health; IRF = inpatient rehabilitation facility; LCI = lower confidence interval; LEJR = lower extremity joint replacement; pp = percentage point; PY = performance year; SNF = skilled nursing facility; UCI = upper confidence interval.

[a] Estimates for HCC score are expressed as units of the score instead of percentage points.

[b] Estimates for the count of HCCs are expressed as numbers instead of percentage points.

Exhibit C-4: Changes in Patient Mix for Patients Who Are Black or African American and Dually Eligible Relative to Patients Who Are Non-Hispanic White and Not Dually Eligible, Elective LEJRs, Baseline Compared With PY 6

Domain	Measure	Change from baseline to PY 6 for CJR vs. change from baseline to PY 6 for control								Difference in relative changes			
		Patients who are Black or African American and dually eligible				Patients who are Non-Hispanic White and non-dually eligible				Estimate of difference (pp)	p-value	90% LCI	90% UCI
		Estimate (pp)	p-value	90% LCI	90% UCI	Estimate (pp)	p-value	90% LCI	90% UCI				
Age	<65	-2.3	0.60	-9.4	4.9	-0.3	0.39	-0.8	0.3	-2.0	0.65	-9.2	5.2
	65-74	7.1	0.03	1.9	12.2	1.1	0.36	-0.9	3.0	6.0	0.09	0.1	11.8
	75-84	-5.4	0.07	-10.3	-0.5	-0.1	0.93	-1.6	1.4	-5.3	0.11	-10.8	0.2
	85+	0.6	0.51	-0.9	2.2	-0.7	0.08	-1.4	0.0	1.3	0.18	-0.3	3.0
Demographics	Female	-1.0	0.71	-5.1	3.2	-0.1	0.85	-1.0	0.8	-0.8	0.74	-5.0	3.3
	Disability, no ESRD	-0.5	0.89	-5.9	5.0	-0.2	0.72	-0.9	0.6	-0.3	0.93	-5.9	5.3
Hierarchical Condition Categories (HCC)	Score [a]	-0.0	0.85	-0.2	0.2	-0.0	0.06	-0.1	-0.0	0.0	0.81	-0.1	0.2
	Count [b]	0.1	0.72	-0.3	0.4	-0.1	0.10	-0.2	-0.0	0.2	0.29	-0.1	0.5
	Count: 0	1.1	0.63	-2.7	4.9	2.0	0.02	0.6	3.4	-0.9	0.64	-4.0	2.2
	Count: 1	1.4	0.56	-2.6	5.4	1.0	0.11	-0.0	2.0	0.5	0.86	-3.7	4.6
	Count: 2	0.8	0.73	-3.0	4.6	-0.4	0.15	-0.9	0.1	1.2	0.61	-2.7	5.2
	Count: 3	0.5	0.85	-4.1	5.2	-1.1	<0.01	-1.7	-0.4	1.6	0.55	-2.8	6.1
	Count: 4+	-3.9	0.28	-9.8	2.1	-1.4	0.22	-3.4	0.5	-2.4	0.49	-8.2	3.3
Health status	Obesity	2.9	0.44	-3.2	9.0	0.2	0.93	-4.0	4.4	2.7	0.47	-3.4	8.8
	Diabetes	-5.3	0.21	-12.2	1.6	0.3	0.69	-0.8	1.3	-5.5	0.15	-11.9	0.8
	Hypertension	-1.8	0.46	-5.8	2.2	0.6	0.54	-1.0	2.3	-2.4	0.30	-6.3	1.4
	Dementia	-1.5	0.30	-3.9	0.9	0.1	0.62	-0.2	0.4	-1.6	0.27	-4.0	0.8
	Congestive heart failure	-0.4	0.91	-6.1	5.4	-0.6	0.31	-1.6	0.4	-0.3	0.95	-5.9	5.3
Prior use	ACH Stay	-0.3	0.90	-3.6	3.1	0.3	0.47	-0.3	0.8	-0.5	0.79	-3.8	2.8
	HH Use	5.3	0.14	-0.5	11.1	0.1	0.89	-0.9	1.1	5.2	0.12	-0.4	10.7
	IRF Stay	-0.8	0.37	-2.4	0.7	-0.2	0.31	-0.4	0.1	-0.7	0.45	-2.2	0.8
	SNF Stay	-1.9	0.15	-4.0	0.3	-0.2	0.53	-0.6	0.3	-1.7	0.23	-4.0	0.6
	Any prior care	7.1	0.05	1.3	13.0	-0.3	0.65	-1.3	0.7	7.4	0.03	1.7	13.1

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 October 2021 that ended by December 2022 (intervention).

Notes: The estimates in this exhibit are the results of unadjusted DiD models on patient characteristics. The estimate of difference is the difference of the unadjusted DiD estimates. Estimates that are significant at the 1%, 5%, or 10% significance level are indicated by red, orange, or yellow shaded cells, respectively. Count characteristics refer to the count of individual HCCs. Any prior care included inpatient hospital, psychiatric hospital, emergency department, SNF, IRF, home health, long-term care hospital, and hospice during the 6 months prior to anchor hospitalization.

ACH = acute care hospital; DDD = triple difference-in-differences; DiD = difference-in-differences; ESRD = end-stage renal disease; HH = home health; IRF = inpatient rehabilitation facility; LCI = lower confidence interval; LEJR = lower extremity joint replacement; pp = percentage point; PY = performance year; SNF = skilled nursing facility; UCI = upper confidence interval.

[a] Estimates for HCC score are expressed as units of the score instead of percentage points.

[b] Estimates for the count of HCCs are expressed as numbers instead of percentage points.

Exhibit C-5: Changes in Patient Mix for Patients Who Are Hispanic Relative to Patients Who Are Non-Hispanic White, Elective LEJRs, Baseline Compared With PY 6

Domain	Measure	Change from baseline to PY 6 for CJR vs. change from baseline to PY 6 for control								Difference in relative changes			
		Patients who are Hispanic				Patients who are Non-Hispanic White				Estimate of difference (pp)	p-value	90% LCI	90% UCI
		Estimate (pp)	p-value	90% LCI	90% UCI	Estimate (pp)	p-value	90% LCI	90% UCI				
Age	<65	3.2	0.07	0.3	6.1	0.1	0.90	-0.9	1.1	3.1	0.08	0.2	6.1
	65-74	-2.3	0.39	-6.7	2.1	1.0	0.45	-1.1	3.1	-3.3	0.23	-7.8	1.2
	75-84	-1.4	0.52	-4.8	2.1	-0.2	0.79	-1.7	1.2	-1.1	0.63	-4.9	2.7
	85+	0.4	0.38	-0.4	1.3	-0.8	0.04	-1.5	-0.2	1.3	0.04	0.3	2.3
Demographics	Female	-0.2	0.93	-4.6	4.1	0.2	0.70	-0.7	1.1	-0.5	0.85	-4.6	3.6
	Eligible for Medicaid	-6.1	0.17	-13.3	1.2	0.2	0.68	-0.7	1.2	-6.3	0.15	-13.5	0.9
	Disability, no ESRD	3.1	0.15	-0.5	6.6	0.2	0.77	-1.0	1.4	2.9	0.18	-0.6	6.4
Hierarchical Condition Categories (HCC)	Score [a]	-0.1	0.14	-0.2	0.0	0.0	0.10	-0.1	0.0	-0.0	0.49	-0.1	0.0
	Count [b]	-0.1	0.16	-0.3	0.0	-0.1	0.11	-0.2	0.0	-0.0	0.80	-0.2	0.1
	Count: 0	2.4	0.03	0.5	4.3	2.0	0.02	0.6	3.4	0.4	0.76	-1.8	2.6
	Count: 1	1.4	0.49	-1.9	4.6	0.9	0.13	-0.1	1.8	0.5	0.81	-2.8	3.8
	Count: 2	0.3	0.86	-2.1	2.6	-0.6	0.07	-1.1	-0.1	0.8	0.61	-1.8	3.4
	Count: 3	-1.2	0.48	-4.0	1.6	-1.0	0.02	-1.6	-0.3	-0.2	0.89	-3.0	2.6
	Count: 4+	-2.8	0.13	-5.9	0.3	-1.3	0.24	-3.3	0.6	-1.5	0.44	-4.7	1.7
Health status	Obesity	-0.7	0.79	-5.3	3.8	0.4	0.88	-3.7	4.4	-1.1	0.59	-4.4	2.2
	Diabetes	0.9	0.75	-3.6	5.4	0.2	0.77	-0.9	1.2	0.7	0.80	-3.8	5.2
	Hypertension	-0.5	0.81	-3.8	2.8	0.7	0.47	-0.9	2.4	-1.2	0.55	-4.6	2.2
	Dementia	-0.4	0.63	-1.8	1.0	-0.0	0.94	-0.3	0.3	-0.4	0.63	-1.8	1.0
	Congestive heart failure	-1.4	0.34	-3.9	1.0	-0.5	0.40	-1.5	0.5	-0.9	0.51	-3.2	1.4

Domain	Measure	Change from baseline to PY 6 for CJR vs. change from baseline to PY 6 for control								Difference in relative changes			
		Patients who are Hispanic				Patients who are Non-Hispanic White				Estimate of difference (pp)	p-value	90% LCI	90% UCI
		Estimate (pp)	p-value	90% LCI	90% UCI	Estimate (pp)	p-value	90% LCI	90% UCI				
Prior use	ACH stay	-0.6	0.39	-1.7	0.5	0.4	0.38	-0.3	1.0	-0.9	0.24	-2.2	0.4
	HH use	-1.7	0.66	-8.1	4.7	0.2	0.79	-0.8	1.2	-1.8	0.61	-7.8	4.1
	IRF stay	-0.0	0.87	-0.4	0.4	-0.2	0.25	-0.5	0.1	0.2	0.59	-0.3	0.6
	SNF stay	-0.4	0.39	-1.0	0.3	-0.2	0.34	-0.6	0.2	-0.1	0.75	-0.8	0.5
	Any prior care	-2.6	0.47	-8.6	3.4	-0.0	0.96	-1.2	1.1	-2.6	0.47	-8.4	3.3

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 October 2021 that ended by December 2022 (intervention).

Notes: The estimates in this exhibit are the results of unadjusted DiD models on patient characteristics. The estimate of difference is the difference of the unadjusted DiD estimates. Estimates that are significant at the 1%, 5%, or 10% significance level are indicated by red, orange, or yellow shaded cells, respectively. Count characteristics refer to the count of individual HCCs. Any prior care included inpatient hospital, psychiatric hospital, emergency department, SNF, IRF, home health, long-term care hospital, and hospice during the 6 months prior to anchor hospitalization.

ACH = acute care hospital; DDD = triple difference-in-differences; DiD = difference-in-differences; ESRD = end-stage renal disease; HH = home health; IRF = inpatient rehabilitation facility; LCI = lower confidence interval; LEJR = lower extremity joint replacement; pp = percentage point; PY = performance year; SNF = skilled nursing facility; UCI = upper confidence interval.

[a] Estimates for HCC score are expressed as units of the score instead of percentage points.

[b] Estimates for the count of HCCs are expressed as numbers instead of percentage points.

Exhibit C-6: Changes in Patient Mix for Patients Who Are Black or African American Relative to Patients Who Are Non-Hispanic White, All LEJRs, Baseline Compared With PY 6

Domain	Measure	Change from baseline to PY 6 for CJR vs. change from baseline to PY 6 for control								Difference in relative changes			
		Patients who are Black or African American				Patients who are Non-Hispanic White				Estimate of difference (pp)	p-value	90% LCI	90% UCI
		Estimate (pp)	p-value	90% LCI	90% UCI	Estimate (pp)	p-value	90% LCI	90% UCI				
Age	<65	2.3	0.33	-1.6	6.3	0.2	0.67	-0.7	1.1	2.1	0.32	-1.4	5.6
	65-74	-1.5	0.38	-4.3	1.3	1.3	0.30	-0.8	3.4	-2.8	0.18	-6.2	0.6
	75-84	-1.3	0.41	-3.8	1.3	0.2	0.81	-1.2	1.6	-1.5	0.46	-4.7	1.8
	85+	0.4	0.59	-0.8	1.7	-1.8	<0.01	-2.8	-0.7	2.2	0.02	0.6	3.7
Demographics	Female	0.6	0.65	-1.5	2.7	-0.2	0.70	-1.2	0.7	0.8	0.53	-1.3	2.8
	Eligible for Medicaid	-1.0	0.72	-5.5	3.6	0.2	0.70	-0.7	1.1	-1.2	0.63	-5.3	2.9
	Disability, no ESRD	3.0	0.09	0.1	5.9	0.4	0.56	-0.8	1.6	2.6	0.07	0.3	4.9
Hierarchical Condition Categories (HCC)	Score [a]	-0.0	0.72	-0.1	0.1	0.0	0.10	-0.1	0.0	-0.0	0.54	-0.0	0.1
	Count [b]	-0.1	0.63	-0.3	0.2	-0.1	0.11	-0.2	0.0	0.1	0.61	-0.1	0.2
	Count: 0	2.2	0.20	-0.6	4.9	2.1	0.01	0.7	3.5	0.1	0.97	-2.1	2.2
	Count: 1	0.4	0.78	-2.1	2.9	1.0	0.11	-0.0	2.0	-0.6	0.68	-2.8	1.7
	Count: 2	-0.7	0.56	-2.9	1.4	-0.4	0.19	-0.8	0.1	-0.4	0.76	-2.4	1.7
	Count: 3	1.7	0.28	-0.9	4.3	-1.0	<0.01	-1.6	-0.4	2.7	0.06	0.4	5.1
	Count: 4+	-3.5	0.13	-7.3	0.3	-1.7	0.18	-3.8	0.4	-1.8	0.35	-5.0	1.4
Health status	Obesity	-1.7	0.50	-6.0	2.5	0.6	0.78	-3.1	4.4	-2.4	0.21	-5.5	0.7
	Diabetes	-3.0	0.20	-6.9	0.8	0.4	0.53	-0.6	1.4	-3.4	0.11	-6.9	0.1
	Hypertension	-0.0	0.99	-2.9	2.9	0.6	0.51	-0.9	2.2	-0.6	0.64	-2.9	1.6
	Dementia	0.2	0.87	-1.7	2.0	-0.3	0.43	-0.8	0.3	0.4	0.66	-1.2	2.1
	Congestive heart failure	2.0	0.19	-0.5	4.4	-0.7	0.29	-1.8	0.4	2.7	0.05	0.4	4.9

Domain	Measure	Change from baseline to PY 6 for CJR vs. change from baseline to PY 6 for control								Difference in relative changes			
		Patients who are Black or African American				Patients who are Non-Hispanic White				Estimate of difference (pp)	p-value	90% LCI	90% UCI
		Estimate (pp)	p-value	90% LCI	90% UCI	Estimate (pp)	p-value	90% LCI	90% UCI				
Prior use	ACH stay	1.3	0.15	-0.2	2.8	0.1	0.70	-0.5	0.8	1.2	0.20	-0.3	2.7
	HH use	4.9	0.05	0.8	8.9	0.0	0.95	-1.0	1.1	4.8	0.02	1.4	8.3
	IRF stay	0.0	0.97	-0.6	0.6	-0.2	0.35	-0.4	0.1	0.2	0.60	-0.4	0.7
	SNF stay	-0.5	0.40	-1.4	0.5	-0.3	0.20	-0.6	0.1	-0.2	0.73	-1.2	0.8
	Any prior care	6.4	<0.01	2.7	10.1	-0.1	0.93	-1.2	1.0	6.4	<0.01	3.0	9.9

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 October 2021 that ended by December 2022 (intervention).

Notes: The estimates in this exhibit are the results of unadjusted DiD models on patient characteristics. The estimate of difference is the difference of the unadjusted DiD estimates. Estimates that are significant at the 1%, 5%, or 10% significance level are indicated by red, orange, or yellow shaded cells, respectively. Count characteristics refer to the count of individual HCCs. Any prior care included inpatient hospital, psychiatric hospital, emergency department, SNF, IRF, home health, long-term care hospital, and hospice during the 6 months prior to anchor hospitalization.

ACH = acute care hospital; DiD = difference-in-differences; ESRD = end-stage renal disease; HH = home health; IRF = inpatient rehabilitation facility; LEJR = lower extremity joint replacement; pp = percentage point; PY = performance year; SNF = skilled nursing facility.

[a] Estimates for HCC score are expressed as units of the score instead of percentage points.

[b] Estimates for the count of HCCs are expressed as numbers instead of percentage points.

Exhibit C-7: Changes in Patient Mix for Patients Who Were Dually Eligible Relative to Patients Who Were Not Dually Eligible, All LEJRs, Baseline Compared With PY 6

Domain	Measure	Change from baseline to PY 6 for CJR vs. change from baseline to PY 6 for control								Difference in relative changes			
		Patients who were dually eligible				Patients who were non-dually eligible				Estimate of difference (pp)	p-value	90% LCI	90% UCI
		Estimate (pp)	p-value	90% LCI	90% UCI	Estimate (pp)	p-value	90% LCI	90% UCI				
Age	<65	2.7	0.14	-0.3	5.8	0.1	0.73	-0.5	0.7	2.6	0.14	-0.3	5.6
	65-74	-1.7	0.26	-4.2	0.8	1.2	0.29	-0.7	3.2	-3.0	0.13	-6.2	0.3
	75-84	-2.0	0.17	-4.4	0.4	0.3	0.77	-1.2	1.7	-2.2	0.24	-5.4	0.9
	85+	1.0	0.38	-0.8	2.8	-1.6	<0.01	-2.6	-0.6	2.6	0.03	0.6	4.6
Demographics	Female	2.6	0.03	0.6	4.6	-0.5	0.42	-1.4	0.5	3.1	0.06	0.4	5.8
	Black or African American	0.9	0.66	-2.4	4.2	0.1	0.80	-0.4	0.6	0.8	0.67	-2.3	3.9
	Hispanic	-3.2	0.22	-7.5	1.1	0.1	0.70	-0.3	0.4	-3.3	0.19	-7.4	0.8
	Disability, no ESRD	0.5	0.80	-2.9	4.0	0.4	0.46	-0.4	1.1	0.2	0.93	-3.2	3.5
Hierarchical Condition Categories (HCC)	Score [a]	0.0	0.93	-0.0	0.0	-0.0	0.12	-0.0	0.0	0.1	0.18	-0.0	0.1
	Count [b]	0.1	0.61	-0.1	0.3	-0.1	0.09	-0.3	0.0	0.2	0.09	0.0	0.4
	Count: 0	0.2	0.82	-1.3	1.7	2.1	0.01	0.7	3.6	-1.9	0.05	-3.5	-0.3
	Count: 1	-0.1	0.93	-1.6	1.5	1.0	0.12	-0.1	2.1	-1.1	0.27	-2.7	0.5
	Count: 2	-0.5	0.57	-1.9	0.9	-0.3	0.33	-0.8	0.2	-0.2	0.84	-1.8	1.4
	Count: 3	-0.5	0.64	-2.5	1.4	-1.1	<0.01	-1.7	-0.4	0.5	0.63	-1.2	2.3
	Count: 4+	0.9	0.65	-2.3	4.2	-1.8	0.16	-4.0	0.3	2.7	0.17	-0.5	6.0
Health status	Obesity	-0.9	0.65	-4.2	2.4	0.4	0.87	-3.5	4.3	-1.3	0.40	-3.8	1.3
	Diabetes	1.0	0.63	-2.4	4.5	0.2	0.74	-0.9	1.4	0.8	0.70	-2.6	4.2
	Hypertension	-0.2	0.85	-2.3	1.9	0.5	0.57	-1.0	2.1	-0.8	0.55	-2.9	1.4
	Dementia	0.8	0.49	-1.1	2.8	-0.1	0.66	-0.6	0.3	1.0	0.39	-0.9	2.8
	Congestive heart failure	1.0	0.34	-0.7	2.8	-0.6	0.33	-1.7	0.4	1.7	0.09	0.1	3.3

Domain	Measure	Change from baseline to PY 6 for CJR vs. change from baseline to PY 6 for control								Difference in relative changes			
		Patients who were dually eligible				Patients who were non-dually eligible				Estimate of difference (pp)	p-value	90% LCI	90% UCI
		Estimate (pp)	p-value	90% LCI	90% UCI	Estimate (pp)	p-value	90% LCI	90% UCI				
Prior use	ACH stay	0.6	0.51	-0.9	2.1	0.2	0.62	-0.4	0.7	0.5	0.59	-0.9	1.8
	HH use	3.1	0.30	-1.8	7.9	0.2	0.79	-0.9	1.2	2.9	0.28	-1.6	7.3
	IRF stay	-0.4	0.33	-1.0	0.3	-0.1	0.42	-0.4	0.1	-0.2	0.43	-0.7	0.3
	SNF stay	-0.5	0.60	-1.9	1.0	-0.3	0.34	-0.7	0.2	-0.2	0.83	-1.8	1.4
	Any prior care	4.5	0.09	0.1	9.0	-0.0	0.96	-1.0	1.0	4.6	0.06	0.5	8.6

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 October 2021 that ended by December 2022 (intervention).

Notes: The estimates in this exhibit are the results of unadjusted DiD models on patient characteristics. The estimate of difference is the difference of the unadjusted DiD estimates. Estimates that are significant at the 1%, 5%, or 10% significance level are indicated by red, orange, or yellow shaded cells, respectively. Count characteristics refer to the count of individual HCCs. Any prior care included inpatient hospital, psychiatric hospital, emergency department, SNF, IRF, home health, long-term care hospital, and hospice during the 6 months prior to anchor hospitalization.

ACH = acute care hospital; DiD = difference-in-differences; ESRD = end-stage renal disease; HH = home health; IRF = inpatient rehabilitation facility; LEJR = lower extremity joint replacement; pp = percentage point; PY = performance year; SNF = skilled nursing facility.

[a] Estimates for HCC score are expressed as units of the score instead of percentage points.

[b] Estimates for the count of HCCs are expressed as numbers instead of percentage points.

Exhibit C-8: Changes in Patient Mix for Patients Who Are Black or African American and Dually Eligible Relative to Patients Who Are Non-Hispanic White and Not Dually Eligible, All LEJRs, Baseline Compared With PY 6

Domain	Measure	Change from baseline to PY 6 for CJR vs. change from baseline to PY 6 for control								Difference in relative changes			
		Patients who are Black or African American and dually eligible				Patients who are Non-Hispanic White and non-dually eligible				Estimate of difference (pp)	p-value	90% LCI	90% UCI
		Estimate (pp)	p-value	90% LCI	90% UCI	Estimate (pp)	p-value	90% LCI	90% UCI				
Age	<65	-3.1	0.43	-9.7	3.5	-0.1	0.62	-0.6	0.3	-3.0	0.46	-9.6	3.6
	65-74	5.5	0.06	0.7	10.2	1.4	0.23	-0.5	3.4	4.0	0.23	-1.5	9.6
	75-84	-4.6	0.06	-8.6	-0.5	0.4	0.68	-1.1	1.8	-5.0	0.08	-9.5	-0.4
	85+	2.2	0.13	-0.2	4.7	-1.7	<0.01	-2.7	-0.7	3.9	0.02	1.2	6.6
Demographics	Female	-0.6	0.81	-4.3	3.2	-0.5	0.37	-1.5	0.4	-0.0	0.99	-3.8	3.8
	Disability, no ESRD	-2.7	0.33	-7.2	1.8	0.0	0.91	-0.7	0.7	-2.7	0.35	-7.4	2.0
Hierarchical Condition Categories (HCC)	Score [a]	-0.1	0.60	-0.2	0.1	-0.1	0.05	-0.1	0.0	0.0	0.96	-0.1	0.2
	Count [b]	0.0	0.81	-0.3	0.4	-0.1	0.08	-0.3	0.0	0.2	0.30	-0.1	0.5
	Count: 0	0.9	0.67	-2.4	4.1	2.1	0.01	0.7	3.5	-1.3	0.42	-3.9	1.4
	Count: 1	0.3	0.89	-3.6	4.2	1.1	0.10	-0.0	2.1	-0.7	0.77	-4.8	3.4
	Count: 2	0.6	0.79	-2.9	4.0	-0.2	0.41	-0.7	0.2	0.8	0.71	-2.8	4.4
	Count: 3	0.8	0.76	-3.7	5.4	-1.2	<0.01	-1.8	-0.6	2.0	0.45	-2.4	6.5
	Count: 4+	-2.6	0.50	-8.8	3.7	-1.8	0.17	-3.9	0.3	-0.8	0.83	-6.9	5.3
Health status	Obesity	1.1	0.75	-4.5	6.6	0.5	0.83	-3.4	4.4	0.6	0.86	-4.8	5.9
	Diabetes	-3.6	0.39	-10.5	3.3	0.4	0.58	-0.7	1.4	-4.0	0.30	-10.2	2.3
	Hypertension	0.2	0.93	-3.6	4.0	0.5	0.56	-1.0	2.1	-0.3	0.87	-3.8	3.1
	Dementia	0.3	0.92	-3.8	4.3	-0.1	0.61	-0.6	0.3	0.4	0.87	-3.6	4.4
	Congestive heart failure	0.8	0.79	-3.9	5.5	-0.8	0.24	-1.8	0.3	1.5	0.59	-3.1	6.2
Prior use	ACH stay	-0.7	0.73	-4.4	2.9	0.1	0.81	-0.4	0.6	-0.8	0.70	-4.4	2.7
	HH use	5.9	0.13	-0.5	12.3	-0.0	0.97	-1.0	1.0	5.9	0.09	0.1	11.7
	IRF stay	-0.9	0.33	-2.5	0.6	-0.1	0.35	-0.4	0.1	-0.8	0.39	-2.3	0.7
	SNF stay	-0.4	0.84	-3.5	2.7	-0.2	0.35	-0.7	0.2	-0.1	0.94	-3.5	3.2
	Any prior care	10.1	<0.01	4.3	15.9	-0.3	0.60	-1.3	0.7	10.4	<0.01	4.8	15.9

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 October 2021 that ended by December 2022 (intervention).

Notes: The estimates in this exhibit are the results of unadjusted DiD models on patient characteristics. The estimate of difference is the difference of the unadjusted DiD estimates. Estimates that are significant at the 1%, 5%, or 10% significance level are indicated by red, orange, or yellow shaded cells, respectively. Count characteristics refer to the count of individual HCCs. Any prior care included inpatient hospital, psychiatric hospital, emergency department, SNF, IRF, home health, long-term care hospital, and hospice during the 6 months prior to anchor hospitalization.

ACH = acute care hospital; DiD = difference-in-differences; ESRD = end-stage renal disease; HH = home health; IRF = inpatient rehabilitation facility; LEJR = lower extremity joint replacement; pp = percentage point; PY = performance year; SNF = skilled nursing facility.

[a] Estimates for HCC score are expressed as units of the score instead of percentage points.

[b] Estimates for the count of HCCs are expressed as numbers instead of percentage points.

Exhibit C-9: Changes in Patient Mix for Patients Who Are Hispanic Relative to Patients Who Are Non-Hispanic White, All LEJRs, Baseline Compared With PY 6

Domain	Measure	Change from baseline to PY 6 for CJR vs. change from baseline to PY 6 for control								Difference in relative changes			
		Patients who are Hispanic				Patients who are Non-Hispanic White				Estimate of difference (pp)	p-value	90% LCI	90% UCI
		Estimate (pp)	p-value	90% LCI	90% UCI	Estimate (pp)	p-value	90% LCI	90% UCI				
Age	<65	3.0	0.07	0.3	5.7	0.2	0.67	-0.7	1.1	2.7	0.10	-0.0	5.5
	65-74	-1.6	0.52	-5.5	2.4	1.3	0.30	-0.8	3.4	-2.9	0.24	-6.9	1.1
	75-84	-0.5	0.77	-3.2	2.3	0.2	0.81	-1.2	1.6	-0.7	0.70	-3.7	2.3
	85+	-0.9	0.25	-2.2	0.4	-1.8	<0.01	-2.8	-0.7	0.8	0.41	-0.8	2.5
Demographics	Female	-0.9	0.73	-5.4	3.5	-0.2	0.70	-1.2	0.7	-0.7	0.77	-4.8	3.4
	Eligible for Medicaid	-4.7	0.28	-11.9	2.5	0.2	0.70	-0.7	1.1	-4.9	0.25	-11.9	2.1
	Disability, No ESRD	3.0	0.13	-0.2	6.2	0.4	0.56	-0.8	1.6	2.6	0.19	-0.7	5.9
Hierarchical Condition Categories (HCC)	Score [a]	-0.1	0.06	-0.2	-0.0	-0.0	0.10	-0.1	0.0	-0.1	0.27	-0.1	0.0
	Count [b]	-0.2	0.04	-0.4	-0.0	-0.1	0.11	-0.2	0.0	-0.1	0.39	-0.2	0.1
	Count: 0	2.3	0.02	0.6	4.0	2.1	0.01	0.7	3.5	0.2	0.87	-1.8	2.2
	Count: 1	1.1	0.59	-2.3	4.5	1.0	0.11	-0.0	2.0	0.1	0.96	-3.4	3.6
	Count: 2	0.8	0.57	-1.5	3.0	-0.4	0.19	-0.8	0.1	1.1	0.46	-1.4	3.6
	Count: 3	-1.3	0.52	-4.6	2.0	-1.0	<0.01	-1.6	-0.4	-0.3	0.89	-3.7	3.1
	Count: 4+	-2.9	0.16	-6.2	0.5	-1.7	0.18	-3.8	0.4	-1.2	0.57	-4.5	2.2
Health status	Obesity	-0.8	0.75	-5.1	3.4	0.6	0.78	-3.1	4.4	-1.5	0.45	-4.6	1.7
	Diabetes	0.3	0.91	-3.8	4.3	0.4	0.53	-0.6	1.4	-0.1	0.96	-4.1	3.9
	Hypertension	-0.7	0.72	-3.8	2.4	0.6	0.51	-0.9	2.2	-1.3	0.50	-4.5	1.9
	Dementia	-1.2	0.21	-2.7	0.4	-0.3	0.43	-0.8	0.3	-0.9	0.31	-2.4	0.6
	Congestive Heart Failure	-1.8	0.16	-3.8	0.3	-0.7	0.29	-1.8	0.4	-1.0	0.35	-2.9	0.8
Prior use	ACH Stay	-1.4	0.06	-2.6	-0.2	0.1	0.70	-0.5	0.8	-1.5	0.06	-2.9	-0.2
	HH Use	-2.4	0.52	-8.7	3.9	0.0	0.95	-1.0	1.1	-2.5	0.48	-8.3	3.3
	IRF Stay	-0.3	0.21	-0.8	0.1	-0.2	0.35	-0.4	0.1	-0.2	0.56	-0.7	0.3
	SNF Stay	-0.7	0.10	-1.4	-0.0	-0.3	0.20	-0.6	0.1	-0.4	0.37	-1.2	0.3
	Any prior care	-3.6	0.29	-9.3	2.1	-0.1	0.93	-1.2	1.0	-3.5	0.28	-9.0	1.9

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 October 2021 that ended by December 2022 (intervention).

Notes: The estimates in this exhibit are the results of unadjusted DiD models on patient characteristics. The estimate of difference is the difference of the unadjusted DiD estimates. Estimates that are significant at the 1%, 5%, or 10% significance level are indicated by red, orange, or yellow shaded cells, respectively. Count characteristics refer to the count of individual HCCs. Any prior care included inpatient hospital, psychiatric hospital, emergency department, SNF, IRF, home health, long-term care hospital, and hospice during the 6 months prior to anchor hospitalization.

ACH = acute care hospital; DiD = difference-in-differences; ESRD = end-stage renal disease; HH = home health; IRF = inpatient rehabilitation facility; LEJR = lower extremity joint replacement; pp = percentage point; PY = performance year; SNF = skilled nursing facility.

[a] Estimates for HCC score are expressed as units of the score instead of percentage points.

[b] Estimates for the count of HCCs are expressed as numbers instead of percentage points.

Appendix D: Claims-Based Impact Analyses – Detailed Tables

A. Parallel Trends

Acronyms	
CJR	Comprehensive Care for Joint Replacement
DiD	difference-in-differences
HH	home health
IRF	inpatient rehabilitation facility
LEJR	lower extremity joint replacement
p	p-value
PY	performance year

One critical assumption of an unbiased difference-in-differences (DiD) estimate was that the treatment and control group outcomes followed parallel trends for the outcome of interest during the baseline period. Another assumption was that these parallel trends would have remained the same in the period when the policy was actually implemented in the absence of policy intervention. While the first assumption could be tested if sufficient baseline data on the Comprehensive Care for Joint Replacement (CJR) and control groups were available, the second assumption was untestable.

We evaluated the parallel trends assumption three ways, with each way testing whether the CJR and control group outcomes followed parallel trends during the baseline period. In the first two methods, we estimated episode-level models for each outcome using baseline data and used both linear and joint F-tests of equality to conclude whether there was evidence to reject the parallel trend assumption. We considered outcomes to fail parallel trends if we rejected the null hypothesis of seemingly parallel trends at the 10% significance level for both tests. In the third method, we descriptively estimated a "Hypothetical DiD," which took into account any potential differential pre-trends between CJR and control. We then compared it with the actual estimated performance year (PY) 6 DiD to gauge how large of an effect potential differential pre-trends would have on the DiD impact estimate.

1. Joint F-Test

For the joint F-test, we report the p-value of an F-test that tested whether the differential between the CJR and control groups was jointly equal across discrete 4-quarter time periods. We included dummy variables for each of the 3 baseline years, interaction terms between the CJR group indicator and each of the year dummies, along with all the risk-adjustment variables that we included in the DiD models, described in the **Appendix B: Data and Methods**.

The joint F-test model was:

$$Y_{i,k,t} = b_0 + b_1 \cdot Year_{1,i} + b_2 \cdot Year_{2,i} + b_3 \cdot Year_{3,i} + b_4 \cdot Year_{1,i} \cdot CJR_k + b_5 \cdot Year_{2,i} \cdot CJR_k + b_6 \cdot Year_{3,i} \cdot CJR_k + X_{i,k}{}' \cdot B + u_{i,k,t}$$

where:

- $Y_{i,k,t}$ was the outcome for the i^{th} episode with an LEJR at hospital k in the baseline period in year t
- $Year_{i,t}$ was an indicator that took the value of 1 if the i^{th} episode was initiated during year t of the baseline period and took the value of 0 otherwise
- $CJR_{i,k}$ was an indicator that took the value of 1 if the i^{th} episode was initiated by a CJR participant hospital k and took the value of 0 otherwise
- $X_{i,k}$ were hospital, geographic, and patient characteristics in the baseline period

and the test was:

$$H_0: b_4 = b_5 = b_6$$

$$H_1: b_4 \neq b_5, or\ b_4 \neq b_6, or\ b_5 \neq b_6$$

2. Linear Test

For the linear test, we report the p-value of a linear slope coefficient of the quarterly difference between the CJR and control group. We included a quarterly indicator; interaction term between the CJR group indicator and the quarterly indicator, along with all the risk-adjustment variables that we included in the DiD models.

The linear test model was:

$$Y_{i,k,t} = b_0 + b_1 \cdot Quarter_{i,t} + b_2 \cdot CJR_k + b_3 \cdot Quarter_{i,t} \cdot CJR_k + {X_{i,k}}' \cdot B + u_{i,k,t}$$

$$Y_{i,k,t} = b_0 + b_1 \cdot Quarter_{i,t} + b_2 \cdot CJR_k + b_3 \cdot Quarter_{i,t} \cdot CJR_k + {X_{i,k}}' \cdot B + u_{i,k,t}$$

where:

- $Y_{i,k,t}$ was the outcome for the i^{th} episode with an LEJR at hospital k in the baseline period in quarter t.
- $Quarter_{i,t}$ was an indicator that took the value of 1 if the i^{th} episode was initiated during quarter t of the baseline period and took the value of 0 otherwise
- $CJR_{i,k}$ was an indicator that took the value of 1 if the i^{th} episode was initiated by a CJR participant hospital k and took the value of 0 otherwise
- $X_{i,k}$ were hospital, geographic, and patient characteristics in the baseline period

and the test was:

$$H_0: b_3 = 0$$

$$H_1: b_3 \neq 0$$

For mandatory CJR hospitals, in the all lower extremity joint replacement (LEJR) sample, home health (HH) payments ($p<0.01$ for the joint test and $p<0.01$ for the linear test), other Part B payments ($p<0.10$ for the joint test and $p<0.10$ for the linear test), 30-day post-episode period (PEP) payments ($p<0.05$ for the joint test and $p<0.05$ for the linear test), first post-acute care (PAC) home with HH ($p<0.01$ for the joint test and $p<0.05$ for the linear test), and any HH use ($p<0.10$ for the joint test and $p<0.05$ for the linear test) failed the parallel trends tests.[120]

For mandatory CJR hospitals, in the elective LEJR sample, HH payments ($p<0.01$ for the joint test and $p<0.01$ for the linear test), other Part B payments ($p<0.10$ for the joint test and $p<0.05$ for the linear test), 30-day PEP payments ($p<0.01$ for the joint test and $p<0.01$ for the linear test), and first PAC home with HH ($p<0.01$ for the joint test and $p<0.05$ for the linear test) failed the parallel trends tests.

[120] See **Appendix B: Data and Methods** for complete definitions of all outcomes, including the first discharge destination outcomes.

For mandatory CJR hospitals, in the fracture LEJR sample, first PAC SNF ($p<0.10$ for the joint test and $p<0.05$ for the linear test) and first PAC institutional rehabilitation facility (IRF) ($p<0.10$ for the joint test and $p<0.10$ for the linear test) failed the parallel trends tests.

Results for both the joint F-Test and linear are shown in **Exhibit D-1**.

Exhibit D-1: Linear and Joint Tests of Parallel Trends for Payment, Utilization, and Quality Metrics, Mandatory CJR Hospitals, Baseline: All LEJR Episodes, Fracture Episodes, and Elective Episodes

Domain	Measure	All LEJR		Elective		Fracture	
		Joint test	Linear test	Joint test	Linear test	Joint test	Linear test
Payments	Episode Payments	p=0.80	p=0.90	p=0.66	p=0.72	p=0.86	p=0.32
	SNF Payments	p=0.75	p=0.73	p=0.70	p=0.88	p=0.41	p=0.28
	IRF Payments	p=0.05	p=0.45	p=0.04	p=0.56	p=0.10	p=0.08
	HH Payments	p<0.01	p<0.01	p<0.01	p<0.01	p=0.39	p=0.46
	Readmission Payments	p=0.31	p=0.28	p=0.23	p=0.15	p=0.90	p=0.71
	Anchor Payments	p<0.01	p=0.14	p<0.01	p=0.11	p=0.92	p=0.64
	Other A Payments	p=0.79	p=0.90	p=0.77	p=0.97	p=0.95	p=0.78
	Other B Payments	p=0.09	p=0.05	p=0.06	p=0.03	p=0.96	p=0.73
	30-day PEP payments	p=0.01	p=0.02	p<0.01	p<0.01	p=0.60	p=0.81
Utilization	First PAC SNF	p=0.17	p=0.19	p=0.14	p=0.28	p=0.06	p=0.04
	First PAC IRF	p=0.05	p=0.45	p=0.03	p=0.61	p=0.06	p=0.07
	First PAC HH	p<0.01	p=0.01	p<0.01	p=0.01	p=0.58	p=0.32
	First PAC home without HH	p=0.41	p=0.31	p=0.41	p=0.33	p=0.49	p=0.84
	SNF days	p=0.14	p=0.06	p=0.16	p=0.11	p=0.38	p=0.13
	IRF days	p=0.36	p=0.34	p=0.41	p=0.40	p=0.86	p=0.58
	HH Visits	p=0.27	p=0.16	p=0.25	p=0.14	p=0.80	p=0.63
	Outpatient PT/OT Visits	p=0.21	p=0.09	p=0.21	p=0.07	p=0.24	p=0.96
	Any HH use	p=0.08	p=0.03	p=0.12	p=0.04	p=0.79	p=0.67
Quality	Unplanned readmission rate	p=0.45	p=0.40	p=0.30	p=0.33	p=0.92	p=0.97
	ED Use rate	p=0.60	p=0.76	p=0.54	p=0.97	p=0.70	p=0.37
	Complication rate	p=0.76	p=0.23	p=0.92	p=0.60	p=0.18	p=0.26
	Mortality rate	p=0.12	p=0.72	p=0.70	p=0.89	p=0.16	p=0.66

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).

***Notes*:** The p-values in this exhibit were the result of risk-adjusted regression models analyzing if the respective CJR and control groups followed parallel trends during the baseline period. For the joint test, we report the p-value of an F-test that tests if the differential between the CJR and control group were jointly equal across annual time periods. For the linear test, we report the p-value of a linear slope coefficient of the quarterly difference between the CJR and control group. Estimates that were significant at the 1%, 5%, or 10% significance level are indicated by red, orange, or yellow shaded cells, respectively. We considered outcomes to *fail* parallel trends if we rejected the null hypothesis of seemingly parallel trends for both tests at the 10% significance level.

ED = emergency department; HH = home health; IRF = inpatient rehabilitation facility; LEJR = lower extremity joint replacement; OT = occupational therapy; NA = not applicable; p = p-value; PAC = post-acute care; PEP = post-episode period; PT = physical therapy; SNF = skilled nursing facility.

3. Hypothetical Difference-in-Differences Method

In addition to using the joint F-Test and linear trend methods to determine whether the CJR and control populations were on differential trends in the baseline period, we conducted a third parallel trend analysis to examine the extent possible deviations in parallel trends could possibly influence our standard PY 6 impact estimates, as reported in **Chapter III: Impact of the Model**. While the calculations of the hypothetical DiD impact estimates were objective, their interpretation was fundamentally subjective. In this section, we describe our approach, our findings, and ultimately our interpretation, with the acknowledgment that the interpretation may differ for different readers.

The first step in calculating the hypothetical impact estimates involved running a differential linear trend parallel trend test between CJR and control using baseline data. The test was almost identical to that described in **Section D.A.2**, with the only difference being that ordinary least squares was used for every outcome.

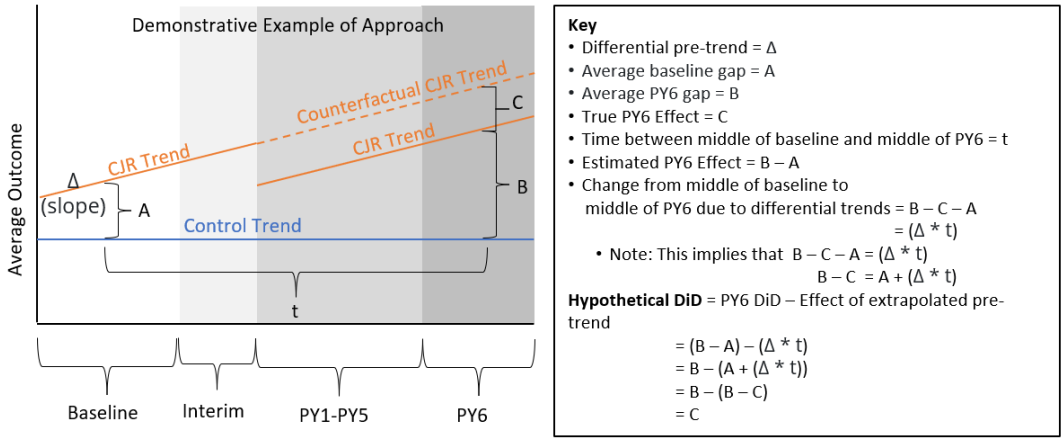
We then used the estimated differential trend (b_3 from the above regression equation in **Section D.A.2**) to calculate the hypothetical DiD using the following equation:

$$\text{Hypothetical DiD} = \text{PY 6 DiD} - (\text{Estimated differential trend} * 34.5),$$

where the PY 6 DiDs were our standard DiD impact estimates (discussed in **Chapter III: Impact of the Model**), the estimated differential trend was the differential trend between CJR and control in the baseline period, and 34.5 represented the number of quarters that were between the middle of the baseline period and the middle of PY 6. Note, due to the nature of the construction, we did not estimate standard errors for the hypothetical DiD estimates.

Exhibit D-2 demonstrates the calculation of the modified DiD algebraically and graphically.

Exhibit D-2: Example of Calculation of Hypothetical DiD



Source: No data were used to make this figure.

Notes: In this illustrating example the slope of the control group trend has been set to zero to allow for a simplified depiction. In reality, the control group trend had a positive or negative slope that varied by outcome. Our calculations appropriately handled these situations.

DiD = difference-in-differences; PY = performance year.

Using **Exhibit D-2** as a **fictional** example, our standard DiD impact estimate was obtained from comparing the average difference between CJR and control in the intervention period (*B*) and the average difference between CJR and control in the baseline period (*A*). However, because CJR and control were on differential trends, the difference between *B* and *A* (*B–A*), would not equal the true treatment effect (*C*). By starting from the standard estimated PY 6 DiD (*B–A*) and subtracting out the deviations in trends that occurred between the middle of the baseline and the middle of PY 6 (the slope of the differential trend multiplied by the time between the baseline and PY 6; $\Delta * t$), we were left with an estimate of the hypothetical CJR effect (*C*).

We interpreted and presented these results holistically. In particular, we paid close attention to both the magnitude and statistical significance of the PY 6 DiD estimate, the magnitude and statistical significance of the differential trend coefficient, and the results for the Joint F-Test and Linear Test (shown in **Exhibit D-1**). While using various pieces of information from multiple results to assist in a single interpretation of an effect of the CJR Model was an inherently subjective process, we have applied the process as consistently across samples and outcomes as possible.

We classified outcomes into four groupings with the following guidelines. Note, while we used these as guidelines, we stress the subjective nature of this analysis, and we present all intermediate pieces of information that were used in our interpretation in this section. For interpretation purposes, we did not think of standard p-value cutoffs, for example, <0.10 as "meaningful" as strictly binding. For example, we would interpret a p-value of 0.12 nearly identical to a p-value of 0.09.

1. Very High Trust:
    a. Outcomes that did not statistically fail the parallel trends tests and where the differences between the baseline trends of the CJR and control samples were of small magnitude.
    b. Outcomes that did not statistically fail the parallel trends tests and had sufficiently large standard errors, such that we did not believe we could extrapolate the trends in any meaningful way.
2. High Trust: Outcomes that did statistically fail the parallel trends tests, but for which the differential trend was small in magnitude. We believe we could interpret the DiD estimates with relative certainty for these outcomes.
3. Low Trust: Outcomes that did not statistically fail the parallel trends test but did have sufficiently large differential baseline trends. For these outcomes, we believe additional caveats may be warranted in the interpretation of the DiD estimate.
4. Very Low Trust: Outcomes that did statistically fail parallel trends tests and had sufficiently large differential pre-trends. We believe strong additional caveats were warranted in the interpretation of the DiD estimates for these outcomes.

Exhibits D-3 though **D-5** show the results of the hypothetical DiD analysis, as well as our interpretation of the degree of trust we had in the DiD point estimate.

Exhibit D-3: Among the All-LEJR Sample, We Had Low Trust Among Some Outcomes, Especially Among the Utilization Outcomes

Domain	Outcome	Differential trend			PY 6 impact			Level of trust in DiD estimate
		Coefficient	P-value	Deviation by PY 6	DiD impact	P-value	Hypothetical DiD	
Payment	Total payments	$3	0.90	$89	-$1,012	0.10	-$1,101	Very High
	SNF payments	-$11	0.56	-$369	-$162	0.60	$208	Very High
	IRF payments	$19	0.42	$651	-$571	0.06	-$1,222	Very High
	HH payments	$17	<0.01	$571	$190	0.34	-$380	Very Low
	Anchor payments	-$4	0.14	-$150	$21	0.53	$171	Very High
	Readmission payments	-$6	0.19	-$194	-$175	0.04	$19	Low
	Other Part A payments	-$0	0.90	-$12	$42	0.25	$54	Very High
	Other Part B payments	-$12	0.05	-$414	-$224	0.10	$189	Very Low
	Post 30 PEP payments	-$8	0.03	-$283	$72	0.04	$355	Very Low
Utilization	First PAC SNF	-0.23	0.16	-7.9	-1.1	0.61	6.8	Low
	First PAC IRF	0.10	0.56	3.6	-3.9	0.05	-7.5	Very High
	First PAC HH	0.32	<0.01	11.2	12.2	0.02	1.0	Very Low
	First PAC home without HH	-0.20	0.19	-6.9	-7.2	0.07	-0.3	Low
	SNF days	0.09	0.06	3.1	0.3	0.70	-2.8	Very Low
	IRF days	-0.01	0.34	-0.4	0.2	0.46	0.6	Very High
	HH Visits	0.04	0.16	1.5	-0.6	0.37	-2.0	High
	Outpatient PT/OT visits	-0.03	0.09	-1.0	0.1	0.84	1.1	High
	Any HH use	0.31	0.03	10.8	8.7	0.17	-2.1	Very Low
Quality	ED use rate	0.01	0.71	0.5	0.4	0.37	-0.2	Very High
	Unplanned readmission rate	-0.03	0.39	-1.1	-0.5	0.12	0.6	Very High
	Complication rate	0.03	0.18	0.9	-0.2	0.30	-1.1	Very High
	Mortality rate	0.01	0.63	0.3	-0.0	0.74	-0.3	Very High

Source: CJR evaluation team analysis of Medicare claims and enrollment data for episodes initiated between January 1, 2012 – December 31, 2014, that ended between April 1, 2012-March 31, 2015 (baseline) and episodes that ended between October 1, 2021 – December 31, 2022 (PY 6 intervention).

***Notes*:** The "Coefficient" column shows the estimated differential trend between CJR and control in the baseline period, the "P-value" shows the associated p-value, and the "Deviation by PY 6" column shows how much trends would have deviated by PY 6, had the estimated differential baseline trend continued from the baseline to PY 6. The "DiD Impact" estimates were reproduced from exhibits in **Appendix D, Section D.2** (below). The "Hypothetical DiD" column takes the difference between the deviation by PY 6 out from the DiD. DiD. Estimates that were significant at the 1%, 5%, or 10% significance level are indicated by red, orange, or yellow shaded cells, respectively. Note, there was no shading for statistical significance for estimate sin the "Hypothetical DiD" column. The "Level of Trust in DiD Estimate" column was subjective and based on the other information in the table.

ED = emergency department; HH = home health; IRF = inpatient rehabilitation facility; LEJR = lower extremity joint replacement; OT = occupational therapy; NA = not applicable; PAC = post-acute care; PEP = post-episode period; PT = physical therapy; SNF = skilled nursing facility.

Exhibit D-4: Among the Elective LEJR Sample, We Had Low Trust Among Some Outcomes, Especially First PAC Discharge Outcomes

Domain	Outcome	Differential trend			PY 6 impact			Level of trust in DiD estimate
		Coefficient	P-value	Deviation by PY 6	DiD impact	P-value	Hypothetical DiD	
Payment	Total payments	-$9	0.72	-$317	-$1,171	0.09	-$854	Very High
	SNF payments	-$6	0.68	-$207	-$333	0.19	-$125	Very High
	IRF payments	$9	0.67	$306	-$410	0.06	-$716	Very High
	HH payments	$18	<0.01	$604	$211	0.33	-$393	Very Low
	Anchor payments	-$5	0.11	-$164	$0	1.00	$163	Low
	Readmission payments	-$9	0.12	-$300	-$164	0.04	$136	Low
	Other Part A payments	$0	0.97	$3	$23	0.32	$20	Very High
	Other Part B payments	-$16	0.03	-$553	-$256	0.08	$297	Very Low
	Post 30 PEP payments	-$9	<0.01	-$312	$21	0.53	$332	Low
Utilization	First PAC SNF	-0.21	0.23	-7.1	-2.2	0.34	4.9	Low
	First PAC IRF	0.04	0.80	1.5	-3.0	0.06	-4.5	Very High
	First PAC HH	0.39	<0.01	13.5	13.7	0.02	0.2	Very Low
	First PAC home without HH	-0.23	0.20	-7.9	-8.5	0.06	-0.6	Low
	SNF days	0.07	0.11	2.5	-0.5	0.58	-2.9	Low
	IRF days	-0.01	0.40	-0.4	0.5	0.08	0.9	Very High
	HH Visits	0.04	0.14	1.5	-0.6	0.37	-2.1	High
	Outpatient PT/OT visits	-0.03	0.07	-1.1	0.1	0.87	1.1	High
	Any HH use	0.34	0.05	11.8	10.3	0.16	-1.5	High
Quality	ED use rate	0.00	0.92	0.1	0.5	0.22	0.4	Very High
	Unplanned readmission rate	-0.04	0.35	-1.2	-0.3	0.22	0.9	Very High
	Complication rate	0.02	0.47	0.5	-0.2	0.16	-0.8	Very High
	Mortality rate	0.00	0.80	0.1	<0.01	0.70	-0.1	Very High

Source: CJR evaluation team analysis of Medicare claims and enrollment data for episodes initiated between January 1, 2012 – December 31, 2014, that ended between April 1, 2012-March 31, 2015 (baseline) and episodes that ended between October 1, 2021 – December 31, 2022 (PY 6 intervention).

***Notes*:** The "Coefficient" column shows the estimated differential trend between CJR and control in the baseline period, the "P-value" shows the associated p-value, and the "Deviation by PY 6" column shows how much trends would have deviated by PY 6, had the estimated differential baseline trend continued from the baseline to PY 6. The "DiD Impact" estimates are reproduced from exhibits in **Appendix D, Section 2** (below). The "Hypothetical DiD" column takes the difference between the deviation by PY 6 out from the DiD. Estimates that were significant at the 1%, 5%, or 10% significance level are indicated by red, orange, or yellow shaded cells, respectively. Note, there was no shading for statistical significance for estimate sin the "Hypothetical DiD" column. The "Level of Trust in DiD Estimate" column was subjective and based on the other information in the table.

ED = emergency department; HH = home health; IRF = inpatient rehabilitation facility; LEJR = lower extremity joint replacement; OT = occupational therapy; NA = not applicable; PAC = post-acute care; PEP = post-episode period; PT = physical therapy; SNF = skilled nursing facility.

Exhibit D-5: Among the Fracture LEJR Sample, We Had Very High Trust in Most Outcomes, Excluding Total, IRF, and SNF Payments

Domain	Outcome	Differential trend			PY 6 impact			Level of trust in DiD estimate
		Coefficient	P-value	Deviation by PY 6	DiD impact	P-value	Hypothetical DiD	
Payment	Total payments	$43	0.32	$1,479	-$354	0.64	-$1,833	Low
	SNF payments	-$61	0.26	-$2,093	$1,575	0.02	$3,668	Low
	IRF payments	$73	0.11	$2,504	-$1,476	0.05	-$3,979	Low
	HH payments	$7	0.38	$251	-$76	0.33	-$327	Very High
	Anchor payments	-$4	0.64	-$124	$121	0.06	$245	Very High
	Readmission payments	$12	0.47	$428	-$179	0.24	-$607	Very High
	Other Part A payments	$3	0.78	$106	$101	0.35	-$6	Very High
	Other Part B payments	$3	0.73	$106	-$52	0.67	-$158	Very High
	Post 30 PEP payments	-$4	0.81	-$125	$351	0.05	$475	Very High
Utilization	First PAC SNF	-0.39	0.04	-13.5	3.1	0.18	16.6	Very Low
	First PAC IRF	0.43	0.07	14.8	-6.4	0.06	-21.2	Very Low
	First PAC HH	-0.06	0.42	-2.1	3.2	0.04	5.3	Very High
	First PAC home without HH	0.02	0.80	0.8	<0.01	0.97	-0.7	Very High
	SNF days	0.15	0.13	5.3	1.7	0.09	-3.6	Very Low
	IRF days	-0.02	0.58	-0.6	-0.1	0.88	0.5	Very High
	HH Visits	0.03	0.63	0.9	-0.2	0.76	-1.1	Very High
	Outpatient PT/OT visits	0.00	0.96	0.1	0.3	0.53	0.3	Very High
	Any HH use	0.08	0.68	2.7	-1.9	0.07	-4.5	Very High
Quality	ED use rate	0.09	0.36	3.1	-0.6	0.49	-3.7	Very High
	Unplanned readmission rate	-0.01	0.95	-0.2	-1.3	0.19	-1.1	Very High
	Complication rate	0.09	0.26	3.1	0.1	0.87	-3.0	Very High
	Mortality rate	0.01	0.63	0.3	-0.1	0.87	-0.4	Very High

Source: CJR evaluation team analysis of Medicare claims and enrollment data for episodes initiated between January 1, 2012 – December 31, 2014, that ended between April 1, 2012-March 31, 2015 (baseline) and episodes that ended between October 1, 2021 – December 31, 2022 (PY 6 intervention).

***Notes*:** The "Coefficient" column shows the estimated differential trend between CJR and control in the baseline period, the "P-value" shows the associated p-value, and the "Deviation by PY 6" column shows how much trends would have deviated by PY 6, had the estimated differential baseline trend continued from the baseline to PY 6. The "DiD Impact" estimates are reproduced from exhibits in **Appendix D, Section 2** (below). The "Hypothetical DiD" column takes the difference between the deviation by PY 6 out from the DiD. Estimates that were significant at the 1%, 5%, or 10% significance level are indicated by red, orange, or yellow shaded cells, respectively. Note, there is no shading for statistical significance for estimate sin the "Hypothetical DiD" column. The "Level of Trust in DiD Estimate" column is subjective and based on the other information in the table.

ED = emergency department; HH = home health; IRF = inpatient rehabilitation facility; LEJR = lower extremity joint replacement; OT = occupational therapy; NA = not applicable; PAC = post-acute care; PEP = post-episode period; PT = physical therapy; SNF = skilled nursing facility.

B. Claims-Based Outcomes

Exhibit D-6: During PY 6, for the All-LEJR Population, Declines in IRF Payments Drove a Reduction in Average Episode Payments

Measure	CJR				Control				DiD	DiD as a percent[b]	p-value	90% LCI	90% UCI
	Baseline episodes (N)	PY 6 episodes (N)	Baseline mean	PY 6 mean	Baseline episodes (N)	PY 6 episodes (N)	Baseline mean	PY 6 mean					
Total payment	114,525	53,328	$28,926	$25,302	141,473	57,426	$28,114	$25,502	-$1,012	-3.5%	0.10	-$2,017	-$7
SNF payment	114,525	53,328	$6,065	$2,974	141,473	57,426	$5,991	$3,062	-$162	-2.7%	0.60	-$662	$339
IRF payment	114,525	53,317	$2,205	$1,203	141,473	57,424	$2,010	$1,578	-$571	-25.9%	0.06	-$1,067	-$74
HH payment*	114,525	53,328	$2,377	$2,078	141,473	57,426	$2,172	$1,683	$190	8.0%	0.34	-$136	$517
Readmission payment	114,525	53,328	$1,179	$990	141,473	57,426	$1,082	$1,068	-$175	-14.8%	0.04	-$316	-$34
Anchor payment	114,525	53,328	$12,153	$12,759	141,473	57,426	$12,134	$12,719	$21	0.2%	0.53	-$34	$76
Other Part A	114,525	53,328	$133	$124	141,473	57,426	$185	$135	$42	31.5%	0.25	-$19	$102
Other Part B*	114,525	53,328	$4,974	$4,556	141,473	57,426	$4,786	$4,592	-$224	-4.5%	0.10	-$450	$1

Source: CJR evaluation team analysis of Medicare claims and enrollment data for episodes initiated between January 1, 2012-December 31, 2014, that ended between April 1, 2012 – March 31, 2015 (baseline) and episodes that ended between October 1, 2021 – December 31, 2022 (PY 6 intervention).

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

Outcome definitions are presented in **Appendix B: Data Sources and Methods**.

CJR = Comprehensive Care for Joint Replacement; HH = home health; IRF = inpatient rehabilitation facility; LCI = lower confidence interval; N = number; PY = performance year; SNF = skilled nursing facility; UCI = upper confidence interval.

[a] Indicates that we believe the CJR and control populations may have been on relatively large differential trends in the baseline period for this outcome.

[b] Percents are calculated by dividing the impact by the risk-adjusted mean value of the outcome for CJR episodes in the baseline period.

Exhibit D-7: During PY 6, for the Elective LEJR Population, Declines in IRF Payments Drove a Reduction in Average Episode Payments

Measure	CJR				Control				DiD	DiD as a percent[b]	p-value	90% LCI	90% UCI
	Baseline episodes (N)	PY 6 episodes (N)	Baseline mean	PY 6 mean	Baseline episodes (N)	PY 6 episodes (N)	Baseline mean	PY 6 mean					
Total payment	96,266	47,070	$25,846	$21,499	122,534	51,142	$24,832	$21,656	-$1,171	-4.5%	0.09	-$2,297	-$44
SNF payment	96,266	47,070	$4,185	$1,306	122,534	51,142	$4,001	$1,454	-$333	-7.9%	0.19	-$754	$89
IRF payment	96,266	47,065	$1,633	$576	122,534	51,141	$1,473	$826	-$410	-25.1%	0.06	-$763	-$58
HH payment*	96,266	47,070	$2,369	$2,012	122,534	51,142	$2,147	$1,580	$211	8.9%	0.33	-$145	$567
Readmission payment	96,266	47,038	$937	$728	122,534	51,116	$845	$799	-$164	-17.5%	0.04	-$296	-$31
Anchor payment	96,266	47,070	$11,978	$12,431	122,534	51,142	$11,941	$12,394	-$0	-0.0%	1.00	-$58	$58
Other Part A	96,266	47,070	$61	$39	122,534	51,142	$85	$40	$23	37.8%	0.32	-$15	$61
Other Part B*	96,266	47,070	$4,766	$4,234	122,534	51,142	$4,545	$4,269	-$256	-5.4%	0.08	-$499	-$13

Source: CJR evaluation team analysis of Medicare claims and enrollment data for episodes initiated between January 1, 2012 – December 31, 2014, that ended between April 1, 2012 – March 31, 2015 (baseline) and episodes that ended between October 1, 2021 – December 31, 2022 (PY 6 intervention).

Notes: The estimates in this exhibit were 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.

Outcome definitions are presented in **Appendix B: Data Sources and Methods**.

CJR = Comprehensive Care for Joint Replacement; HH = home health; IRF = inpatient rehabilitation facility; LCI = lower confidence interval; N = number; PY = performance year; SNF = skilled nursing facility; UCI = upper confidence interval.

[a] Indicates that we believe the CJR and control populations may have been on relatively large differential trends in the baseline period for this outcome.

[b] Percents are calculated by dividing the impact by the risk-adjusted mean value of the outcome for CJR episodes in the baseline period.

Exhibit D-8: During PY 6, for the Fracture LEJR Population, There Were Large, Offsetting Changes in IRF and SNF Payments, Resulting in No Change for Average Episode Payments

Measure	CJR				Control				DiD	DiD as a percent[b]	p-value	90% LCI	90% UCI
	Baseline episodes (N)	PY 6 episodes (N)	Baseline mean	PY 6 mean	Baseline episodes (N)	PY 6 episodes (N)	Baseline mean	PY 6 mean					
Total payment*	18,259	6,258	$47,101	$48,257	18,939	6,284	$47,094	$48,604	-$354	-0.8%	0.64	-$1,623	$915
SNF payment*	18,259	6,252	$16,711	$15,258	18,939	6,283	$17,389	$14,360	$1,575	9.4%	0.02	$480	$2,671
IRF payment*	18,026	6,184	$5,414	$5,909	18,822	6,254	$4,932	$6,904	-$1,476	-27.3%	0.05	-$2,721	-$231
HH payment	18,259	6,258	$2,437	$2,501	18,939	6,284	$2,335	$2,475	-$76	-3.1%	0.33	-$204	$52
Readmission payment	18,259	6,258	$2,564	$2,630	18,939	6,284	$2,485	$2,730	-$179	-7.0%	0.24	-$430	$72
Anchor payment	18,259	6,258	$13,212	$14,827	18,939	6,284	$13,247	$14,742	$121	0.9%	0.06	$16	$225
Other Part A	18,259	6,258	$573	$592	18,939	6,284	$736	$655	$101	17.6%	0.35	-$78	$279
Other Part B	18,259	6,258	$6,226	$6,528	18,939	6,284	$6,184	$6,537	-$52	-0.8%	0.67	-$253	$149

Source: CJR evaluation team analysis of Medicare claims and enrollment data for episodes initiated between January 1, 2012 – December 31, 2014, that ended between April 1, 2012 – March 31, 2015 (baseline) and episodes that ended between October 1, 2021 – December 31, 2022 (PY 6 intervention).

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

Outcome definitions are presented in **Appendix B: Data Sources and Methods**.

CJR = Comprehensive Care for Joint Replacement; HH = home health; IRF = inpatient rehabilitation facility; LCI = lower confidence interval; N = number; PY = performance year; SNF = skilled nursing facility; UCI = upper confidence interval.

[a] Indicates that we believe the CJR and control populations may have been on relatively large differential trends in the baseline period for this outcome.

[b] Percents are calculated by dividing the impact by the risk-adjusted mean value of the outcome for CJR episodes in the baseline period.

Exhibit D-9: During PY 6, for the All-LEJR Population, the CJR Model Led to a Large Shift Away From Patients First Being Discharged to IRFs

Measure	CJR				Control				DiD	DiD as a percent[b]	p-value	90% LCI	90% UCI
	Baseline episodes (N)	PY 6 episodes (N)	Baseline mean	PY 6 mean	Baseline episodes (N)	PY 6 episodes (N)	Baseline mean	PY 6 mean					
First PAC SNF*	114,525	53,327	42.0%	15.6%	141,473	57,426	40.9%	15.7%	-1.1	-2.6%	0.61	-4.7	2.5
First PAC IRF	114,525	53,327	13.7%	5.0%	141,473	57,426	12.6%	7.8%	-3.9	-28.3%	0.05	-7.2	-0.6
First PAC HH*	114,525	53,327	35.6%	55.7%	141,473	57,426	33.2%	41.1%	12.2	34.3%	0.02	3.5	20.9
First PAC home without HH*	114,525	53,327	8.8%	23.7%	141,473	57,426	13.3%	35.4%	-7.2	-81.7%	0.07	-13.8	-0.7
SNF days*	51,542	8,601	26.7	23.0	58,508	8,015	27.2	23.1	0.3	1.2%	0.70	-1.1	1.7
IRF days	16,415	2,883	11.5	12.2	17,844	3,900	11.4	11.8	0.2	2.0%	0.46	-0.3	0.8
HH visits	80,813	37,475	16.8	13.8	99,158	33,483	16.3	13.9	-0.6	-3.4%	0.37	-1.6	0.5
Any HH use*	114,525	53,328	73.3%	72.4%	141,473	57,426	69.0%	59.3%	8.7	11.9%	0.17	-1.8	19.2
Outpatient PT/OT visits	69,289	44,696	12.98	14.32	90,157	49,041	13.30	14.53	0.1	0.8%	0.84	-0.8	1.0

Source: CJR evaluation team analysis of Medicare claims and enrollment data for episodes initiated between January 1, 2012 – December 31, 2014, that ended between April 1, 2012 – March 31, 2015 (baseline) and episodes that ended between October 1, 2021 – December 31, 2022 (PY 6 intervention).

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

Outcome definitions are presented in **Appendix B: Data Sources and Methods**.

CJR = Comprehensive Care for Joint Replacement; HH = home health; IRF = inpatient rehabilitation facility; LCI = lower confidence interval; LOS = length of stay; N = number; PAC = post-acute care; PTOT = physical therapy and occupational therapy; PY = performance year; SNF = skilled nursing facility; UCI = upper confidence interval.

[a] Indicates that we believe the CJR and control populations may have been on relatively large differential trends in the baseline period for this outcome.

[b] Percents are calculated by dividing the impact by the risk-adjusted mean value of the outcome for CJR episodes in the baseline period.

Exhibit D-10: During PY 6, for the Elective LEJR Population, the CJR Model Led to a Large Shift Away From Patients First Being Discharged to IRFs

Measure	CJR				Control				DiD	DiD as a percent[b]	p-value	90% LCI	90% UCI
	Baseline episodes (N)	PY 6 episodes (N)	Baseline mean	PY 6 mean	Baseline episodes (N)	PY 6 episodes (N)	Baseline mean	PY 6 mean					
First PAC SNF*	96,266	47,070l	38.2%	8.7%	122,534	51,142	36.4%	9.1%	-2.2	-5.9%	0.34	-6.1	1.6
First PAC IRF	96,266	47,070	11.4%	2.1%	122,534	51,142	10.7%	4.3%	-3.0	-26.2%	0.06	-5.6	-0.4
First PAC HH*	96,266	47,070	40.7%	62.5%	122,534	51,142	38.0%	46.1%	13.7	33.7%	0.02	4.3	23.1
First PAC home without HH*	96,266	47,070	9.7%	26.8%	122,534	51,142	15.0%	40.5%	-8.5	-87.3%	0.06	-16.0	-1.0
SNF days*	38,172	4,623	20.1	18.1	45,034	4,391	20.4	18.8	-0.5	-2.3%	0.58	-1.9	0.9
IRF days	11,325	1,356	10.2	11.6	12,356	1,986	10.1	11.1	0.5	4.7%	0.08	0.0	0.9
HH visits	68,802	33,194	16.2	12.9	87,138	29,160	15.6	13.0	-0.6	-3.9%	0.37	-1.8	0.5
Any HH use	96,266	47,070	74.0%	73.0%	122,534	51,142	69.4%	58.1%	10.3	13.9%	0.16	-1.7	22.2
Outpatient PT/OT visits	64,933	42,577	13.14	14.43	85,782	46,994	13.45	14.66	0.1	0.6%	0.87	-0.8	0.9

Source: CJR evaluation team analysis of Medicare claims and enrollment data for episodes initiated between January 1, 2012 – December 31, 2014, that ended between April 1, 2012 – March 31, 2015 (baseline) and episodes that ended between October 1, 2021 – December 31, 2022 (PY 6 intervention).

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

Outcome definitions are presented in **Appendix B: Data Sources and Methods**.

CJR = Comprehensive Care for Joint Replacement; HH = home health; IRF = inpatient rehabilitation facility; LCI = lower confidence interval; LOS = length of stay; N = number; PAC = post-acute care; PTOT = physical therapy and occupational therapy; PY = performance year; SNF = skilled nursing facility; UCI = upper confidence interval.

[a] Indicates that we believe the CJR and control populations may have been on relatively large differential trends in the baseline period for this outcome.

[b] Percents are calculated by dividing the impact by the risk-adjusted mean value of the outcome for CJR episodes in the baseline period.

Exhibit D-11: During PY 6, for the Fracture LEJR Population, the CJR Model Led to a Large Shift Toward Patients First Being Discharged Home with HH

Measure	CJR				Control				DiD	DiD as a percent[b]	p-value	90% LCI	90% UCI
	Baseline episodes (N)	PY 6 episodes (N)	Baseline mean	PY 6 mean	Baseline episodes (N)	PY 6 episodes (N)	Baseline mean	PY 6 mean					
First PAC SNF*	18,259	6,257	64.4%	58.1%	18,939	6,284	65.3%	55.9%	3.1	4.9%	0.18	-0.7	7.0
First PAC IRF*	18,259	6,257	27.1%	22.5%	18,939	6,284	25.4%	27.2%	-6.4	-23.5%	0.06	-12.0	-0.7
First PAC HH	18,259	6,257	5.2%	14.9%	18,939	6,284	5.6%	12.1%	3.2	60.9%	0.04	0.6	5.8
First PAC home without HH	18,259	6,257	3.2%	4.4%	18,939	6,284	3.6%	4.8%	0.0	0.9%	0.97	-1.1	1.1
SNF days*	13,370	3,978	43.1	37.0	13,474	3,624	44.3	36.5	1.7	3.9%	0.09	0.0	3.4
IRF days	5,090	1,527	14.1	14.0	5,488	1,914	13.9	13.8	-0.1	-0.4%	0.88	-0.7	0.5
HH visits	12,011	4,281	21.0	19.9	12,020	4,323	20.5	19.6	-0.2	-0.8%	0.76	-1.1	0.7
Any HH use	18,259	6,258	69.3%	69.9%	18,939	6,284	67.5%	70.0%	-1.9	-2.7%	0.07	-3.6	-0.2
Outpatient PT/OT visits	4,356	2,119	10.58	12.32	4,375	2,047	10.72	12.13	0.3	3.1%	0.53	-0.5	1.2

Source: CJR evaluation team analysis of Medicare claims and enrollment data for episodes initiated between January 1, 2012 – December 31, 2014, that ended between April 1, 2012 – March 31, 2015 (baseline) and episodes that ended between October 1, 2021 – December 31, 2022 (PY 6 intervention).

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

Outcome definitions are presented in **Appendix B: Data Sources and Methods**.

CJR = Comprehensive Care for Joint Replacement; HH = home health; IRF = inpatient rehabilitation facility; LCI = lower confidence interval; LOS = length of stay; N = number; PAC = post-acute care; PTOT = physical therapy and occupational therapy; PY = performance year; SNF = skilled nursing facility; UCI = upper confidence interval.

[a] Indicates that we believe the CJR and control populations may have been on relatively large differential trends in the baseline period for this outcome.

[b] Percents are calculated by dividing the impact by the risk-adjusted mean value of the outcome for CJR episodes in the baseline period.

Exhibit D-12: During PY 6, for the All-LEJR Population, Quality of Care Did Not Change

Measure	CJR				Control				DiD	DiD as a percent[a]	p-value	90% LCI	90% UCI
	Baseline episodes (N)	PY 6 episodes (N)	Baseline mean	PY 6 mean	Baseline episodes (N)	PY 6 episodes (N)	Baseline mean	PY 6 mean					
Unplanned readmission rate	114,500	53,314	10.0%	6.9%	141,450	57,417	9.6%	6.9%	-0.5	-4.6%	0.12	-0.9	0.0
ED use rate	114,500	53,314	12.9%	13.2%	141,450	57,417	12.4%	12.4%	0.4	2.8%	0.37	-0.3	1.0
Mortality rate	117,415	54,253	2.5%	2.4%	144,680	58,347	2.6%	2.5%	-0.0	-1.8%	0.74	-0.3	0.2
Complication rate	114,500	53,303	4.0%	2.4%	141,450	57,415	3.8%	2.4%	-0.2	-4.8%	0.30	-0.5	0.1

Source: CJR evaluation team analysis of Medicare claims and enrollment data for episodes initiated between January 1, 2012 – December 31, 2014, that ended between April 1, 2012 – March 31, 2015 (baseline) and episodes that ended between October 1, 2021 – December 31, 2022 (PY 6 intervention).

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

Outcome definitions are presented in **Appendix B: Data Sources and Methods**.

CJR = Comprehensive Care for Joint Replacement; ED = emergency department; LCI = lower confidence interval; N = number; PY = performance year; UCI = upper confidence interval.

[a] Percents are calculated by dividing the impact by the risk-adjusted mean value of the outcome for CJR episodes in the baseline period.

Exhibit D-13: During PY 6, for the Elective LEJR Population, Quality of Care Did Not Change

Measure	CJR				Control				DiD	DiD as a percent[a]	p-value	90% LCI	90% UCI
	Baseline episodes (N)	PY 6 episodes (N)	Baseline mean	PY 6 mean	Baseline episodes (N)	PY 6 episodes (N)	Baseline mean	PY 6 mean					
Unplanned readmission rate	96,244	47,031	8.0%	5.4%	122,516	51,109	7.7%	5.4%	-0.3	-3.8%	0.22	-0.7	0.1
ED use rate	96,244	47,058	12.0%	12.3%	122,516	51,134	11.6%	11.3%	0.5	4.5%	0.22	-0.2	1.3
Mortality rate	96,783	47,183	0.5%	0.5%	123,165	51,279	0.5%	0.5%	-0.0	-4.3%	0.70	-0.1	0.1
Complication rate	96,244	47,026	3.0%	1.57%	122,516	51,108	2.9%	1.8%	-0.2	-7.3%	0.16	-0.5	0.0

Source: CJR evaluation team analysis of Medicare claims and enrollment data for episodes initiated between January 1, 2012 – December 31, 2014, that ended between April 1, 2012 – March 31, 2015 (baseline) and episodes that ended between October 1, 2021 – December 31, 2022 (PY 6 intervention).

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

Outcome definitions are presented in **Appendix B: Data Sources and Methods**.

CJR = Comprehensive Care for Joint Replacement; ED = emergency department; LCI = lower confidence interval; N = number; PY = performance year; UCI = upper confidence interval.

[a] Percents are calculated by dividing the impact by the risk-adjusted mean value of the outcome for CJR episodes in the baseline period.

Exhibit D-14: During PY 6, for the Fracture LEJR Population, Quality of Care Did Not Change

Measure	CJR				Control				DiD	DiD as a percent[a]	p-value	90% LCI	90% UCI
	Baseline episodes (N)	PY 6 episodes (N)	Baseline mean	PY 6 mean	Baseline episodes (N)	PY 6 episodes (N)	Baseline mean	PY 6 mean					
Unplanned readmission rate	18,256	6,251	21.1%	15.4%	18,934	6,282	10.72	12.13	-1.3	-6.1%	0.19	-2.9	0.3
ED use rate	18,256	6,251	18.0%	18.5%	18,934	6,282	20.5%	16.1%	-0.6	-3.4%	0.49	-2.0	0.8
Mortality rate	20,632	7,038	12.9%	12.3%	21,515	7,042	17.7%	18.8%	-0.1	-0.9%	0.87	-1.3	1.0
Complication rate	18,256	6,245	9.8%	6.4%	18,934	6,281	9.6%	6.1%	0.1	1.0%	0.87	-0.9	1.1

Source: CJR evaluation team analysis of Medicare claims and enrollment data for episodes initiated between January 1, 2012 – December 31, 2014, that ended between April 1, 2012 – March 31, 2015 (baseline) and episodes that ended between October 1, 2021 – December 31, 2022 (PY 6 intervention).

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

Outcome definitions are presented in **Appendix B: Data and Methods**.

CJR = Comprehensive Care for Joint Replacement; ED = emergency department; LCI = lower confidence interval; N = number; PY = performance year; UCI = upper confidence interval.

[a] Percents are calculated by dividing the impact by the risk-adjusted mean value of the outcome for CJR episodes in the baseline period.

Exhibit D-15: 30-Day Post-Episode Payments for All LEJR, Elective Only, and Hip Fracture Episodes

Measure	CJR				Control				DiD	DiD as a percent[a]	p-value	90% LCI	90% UCI
	Baseline episodes (N)	PY 6 episodes (N)	Baseline mean	PY 6 mean	Baseline episodes (N)	PY 6 episodes (N)	Baseline mean	PY 6 mean					
All LEJR episodes*	114,525	53,328	$1,448	$1,759	141,473	57,426	$1,482	$1,721	$72	5.0%	0.04	$14	$129
Elective LEJR episodes*	96,266	47,065	$1,123	$1,415	122,534	51,141	$1,122	$1,393	$21	1.8%	0.53	-$34	$75
Hip fracture LEJR episodes	18,259	6,252	$3,368	$3,699	18,939	6,276	$3,612	$3,592	$351	10.4%	0.05	$51	$650

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 ending between October 1, 2021, and December 31, 2022 (PY 6 intervention).

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

Outcome definitions are presented in **Appendix B: Data and Methods**.

CJR = Comprehensive Care for Joint Replacement; DiD = difference-in-differences; LCI = lower confidence interval; LEJR = lower-extremely joint replacement; N = number; PY = performance year; UCI = upper confidence interval.

* Indicates that we believed the CJR and control populations may have been on meaningfully differential trends in the baseline period for this outcome.

Exhibit D-16: Volume Results for Elective LEJR

Population	CJR			Control			Impact (DiD)	p-value	90% LCI	90% UCI
	PY 6 bene-years (N)	Baseline mean	PY 6 mean	PY 6 Bene-years (N)	Baseline mean	PY 6 mean				
All beneficiaries	4,080,740	1,052	1,050	3,937,453	1,110	1,061	47	0.10	0	94

Source: CJR evaluation team analysis of Medicare claims and enrollment data for Medicare FFS beneficiaries enrolled in Medicare between 2012 and 2015 (baseline) or during 2022 (intervention).

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

Outcome definitions are presented in **Appendix B: Data Sources and Methods**.

CJR = Comprehensive Care for Joint Replacement; DiD = difference-in-differences; LCI = lower confidence interval; LEJR = lower-extremely joint replacement; N = number; PY = performance year; UCI = upper confidence interval.

C. Sensitivity Analysis

In this section, we describe the results of the sensitivity analyses we conducted to examine the robustness of the reported impact estimates. Specifically, we examined our model's robustness to control for Coronavirus Disease 2019 (COVID-19) episodes. Previous annual reports have conducted numerous other sensitivity analyses and should be consulted by readers looking for further confirmation of methodological robustness.

Our sensitivity test dropped all episodes with a COVID diagnosis, which made up about 1.7% of our sample. Since payments for CJR episodes with a COVID diagnosis were capped at the target price during the Public Health Emergency,[121] this could have influenced hospital incentives and behavior. While our main analysis included controls for COVID episodes, it was possible that merely controlling was inadequate for dealing with the effects of capped target prices on model outcomes.

We did not find a meaningful difference between our main analysis and our COVID sensitivity test.

The two rows of **Exhibit D-17** show the results of our main analysis and our COVID sensitivity test. Comparing our main estimate for total episode payments in the full sample to the same estimate with dropped COVID episodes, we could see there was only a $26 difference in the point estimate, which was not large enough to be a statistically significant difference at the 10% level. Given that the difference between the episodes was neither statistically nor substantially different, we concluded that our main approach to controlling for COVID episode target price capping was adequate.

Exhibit D-17: Sensitivity Analysis Results

Outcome	Analysis	DiD	p-value	90% LCI	90% UCI	Total N
Episode payments	Main estimate	-$1,012	0.09	-$2,003	-$21	366,752
	Drop COVID episodes	-$1,028	0.10	-$2,046	-$10	360,667

Source: CJR evaluation team analysis of Medicare claims and enrollment data for episodes initiated between January 1, 2012 – December 31, 2014, that ended between April 1, 2012 – March 31, 2015 (baseline) and episodes that ended between October 1, 2021 – December 31, 2022 (PY 6 intervention).

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

LCI = lower confidence interval; N = number; UCI = upper confidence interval.

[121] The COVID-19 Public Health Emergency expired on May 11, 2023, so it covered all of PY 6, which ended on December 31, 2022. Source: https://www.hhs.gov/about/news/2023/05/09/fact-sheet-end-of-the-covid-19-public-health-emergency.html

Appendix E: Patient Survey Results

A. Comparing Recalled Functional Status in the Week Prior to Hospitalization Between CJR and Control Respondents

Exhibit E-1: Summary Statistics in Prehospital Functional Status Between CJR and Control Respondents Across All LEJRs and for Hip Fractures

Measure	Response range	All LEJRs			Hip fractures		
		CJR mean	Control mean	Std. diff.	CJR mean	Control mean	Std. diff.
Walking without rest	-4 to 4	2.8	2.7	<0.1	3.3	3.2	<0.1
Going up or down stairs	-3 to 3	2.3	2.2	<0.1	3.0	2.9	0.2
Rising from sitting	-4 to 4	2.9	2.9	<0.1	4.0	4.0	0.1
Standing	-4 to 4	3.1	3.1	<0.1	4.1	4.2	0.1
Use of a mobility device	-2 to 2	2.3	2.3	<0.1	2.4	2.4	<0.1
Getting on or off the toilet	-4 to 4	3.2	3.1	<0.1	4.2	4.2	0.1
Pain limiting regular activities	-4 to 4	2.1	2.1	<0.1	4.3	4.4	0.2
Medication intensity	-3 to 3	2.9	2.8	<0.1	3.7	3.7	<0.1

Source: CJR evaluation team analysis of patient survey data for LEJR episodes with inpatient discharge or outpatient procedure in July-October 2022.

***Notes*:** Means and standardized differences are unweighted. LEJR = lower extremity join replacement; Std. diff. = standardized difference.

Exhibit E-2: Summary Statistics in Prehospital Functional Status Between CJR and Control Respondents for Elective Inpatient and Elective Outpatient Episodes

Measure	Response range	Elective inpatient			Elective outpatient		
		CJR mean	Control mean	Std. diff.	CJR mean	Control mean	Std. diff.
Walking without rest	-4 to 4	2.6	2.5	<0.1	2.9	2.8	0.1
Going up or down stairs	-3 to 3	2.1	2.0	0.1	2.2	2.2	0.1
Rising from sitting	-4 to 4	2.7	2.7	<0.1	2.8	2.7	0.0
Standing	-4 to 4	2.9	2.9	<0.1	3.1	3.0	0.1
Use of a mobility device	-2 to 2	2.2	2.1	<0.1	2.4	2.3	0.0
Getting on or off the toilet	-4 to 4	3.0	2.9	<0.1	3.1	3.0	0.1
Pain limiting regular activities	-4 to 4	1.8	1.7	<0.1	1.9	1.8	0.1
Medication intensity	-3 to 3	2.7	2.6	0.1	2.8	2.7	0.1

Source: CJR evaluation team analysis of patient survey data for LEJR episodes with inpatient discharge or outpatient procedure in July-October 2022.

***Notes*:** Means and standardized differences are unweighted.

CJR = Comprehensive Care for Joint Replacement; Std. diff. = standardized difference.

B. Comparing Claims-Based Patient Characteristics Between CJR and Control Respondents

Exhibit E-3: Summary Statistics in Patient Characteristics for CJR and Control Respondents Across All LEJRs and for Hip Fractures

Measure		All LEJRs			Hip fractures		
		CJR mean	Control mean	Std. diff.	CJR mean	Control mean	Std. diff.
Claims covariates	Hip fracture	12.0%	14.1%	0.1	-	-	-
	Knee procedure	55.1%	53.8%	<0.1	-	-	-
	Female	64.9%	65.3%	<0.1	70.2%	74.0%	0.1
	Age	75.2	75.2	<0.1	81.2	81.2	<0.1
	MCC	3.6%	5.0%	0.1	17.1%	24.8%	0.2
	Eligible for Medicaid	4.4%	3.9%	<0.1	6.9%	4.8%	0.1
	Disability, no ESRD	7.6%	8.4%	<0.1	7.6%	6.5%	<0.1
	Assignment to ACO	53.6%	55.5%	<0.1	47.4%	53.0%	0.1
	Prior SNF stay[a]	2.0%	1.6%	<0.1	5.8%	4.8%	<0.1
	HCC score	1.4	1.4	<0.1	2.4	2.3	0.1
Survey covariates	Prehospital functional status	47.8%	46.9%	<0.1	70.9%	71.7%	<0.1
	Proxy response	2.4%	2.6%	<0.1	12.0%	11.9%	<0.1
	Asian, Pacific Islander, American Indian, or Alaska Native	3.0%	1.9%	0.1	4.1%	2.2%	0.2
	Black or African American	3.8%	4.5%		2.5%	2.3%	
	Hispanic	2.8%	2.7%		3.1%	*	
	Non-Hispanic White	90.4%	90.8%		90.3%	94.2%	
	Race/ethnicity non-response	10.8%	10.2%	<0.1	12.4%	9.6%	0.1
	Some high school	2.5%	2.8%	0.1	6.4%	3.2%	0.2
	High school or GED	23.3%	25.6%		29.2%	32.6%	
	College–less than bachelor's degree	31.5%	31.9%		28.3%	27.4%	
	College–bachelor's degree	16.4%	18.1%		12.9%	18.2%	
	More than bachelor's degree	26.2%	21.6%		23.2%	18.6%	
	Education non-response	5.7%	5.6%	<0.1	9.6%	7.4%	0.1

Source: CJR evaluation team analysis of patient survey data for LEJR episodes with inpatient discharge or outpatient procedure in July – October 2022.

***Notes*:** Means and standardized differences are unweighted. MCC = Major complication or comorbidity (DRG 469 for elective, DRG 521 for hip fracture).

ACO = Accountable Care Organization; ESRD = end-stage renal disease; GED = general educational development; HCC = hierarchical condition category; LEJR = lower extremity joint replacement; MS-DRG = Medicare Severity-Diagnosis Related Group; SNF = skilled nursing facility; * = Results suppressed due to small sample sizes.

[a] Stay in skilled nursing facility or nursing home in six months prior to admission.

Exhibit E-4: Summary Statistics in Patient Characteristics for CJR and Control Respondents With Elective Inpatient or Elective Outpatient Episodes

Measure		Elective inpatient			Elective outpatient		
		CJR mean	Control mean	Std. diff.	CJR mean	Control mean	Std. diff.
Claims covariates	Hip fracture	-	-	-	-	-	-
	Knee procedure	62.4%	62.5%	<0.1	62.7%	62.7%	<0.1
	Female	66.8%	68.9%	<0.1	62.3%	60.5%	<0.1
	Age	75.3	75.0	<0.1	73.7	73.7	<0.1
	MCC	4.3%	4.2%	<0.1	-	-	-
	Eligible for Medicaid	5.3%	4.3%	<0.1	3.2%	3.4%	<0.1
	Disability, no ESRD	8.7%	9.7%	<0.1	6.8%	8.1%	<0.1
	Assignment to ACO	53.7%	56.8%	0.1	55.0%	55.3%	<0.1
	Prior SNF stay[a]	2.8%	2.4%	<0.1	*	*	0.1
	HCC score	1.5	1.5	<0.1	1.2	1.2	<0.1
Survey covariates	Prehospital functional status	42.5%	40.6%	0.1	46.2%	44.5%	0.1
	Proxy response	1.6%	1.5%	<0.1	0.6%	0.8%	<0.1
	Asian, Pacific Islander, American Indian, or Alaska Native	3.2%	1.7%	0.1	2.6%	2.0%	0.1
	Black or African American	5.1%	5.4%		3.2%	4.6%	
	Hispanic	3.1%	2.3%		2.6%	3.4%	
	Non-Hispanic White	88.7%	90.6%		91.6%	90.0%	
	Race/ethnicity non-response	11.4%	9.9%	<0.1	10.1%	10.5%	<0.1
	Some high school	2.4%	3.2%	0.1	1.7%	2.4%	0.1
	High school or GED	24.4%	26.4%		21.1%	23.2%	
	College–less than bachelor's degree	31.9%	31.7%		32.0%	33.3%	
	College–bachelor's degree	15.6%	17.6%		17.8%	18.4%	
	More than bachelor's degree	25.6%	21.0%		27.4%	22.8%	
	Education non-response	5.8%	5.1%	<0.1	4.7%	5.4%	<0.1

Source: CJR evaluation team analysis of patient survey data for LEJR episodes with inpatient discharge or outpatient procedure in July-October 2022.

***Notes*:** Means and standardized differences are unweighted. MCC = Major complication or comorbidity (DRG 469 for elective, DRG 521 for hip fracture).

ACO = Accountable Care Organization; ESRD = end-stage renal disease; GED = general education development; HCC = hierarchical condition category; LEJR = lower extremity joint replacement; MS-DRG = Medicare Severity-Diagnosis Related Group; SNF = skilled nursing facility; * = Results suppressed due to small sample sizes.

[a] Stay in skilled nursing facility or nursing home in six months prior to admission.

C. Additional Results

Exhibit E-5: Risk-Adjusted Survey-Based Results for Change in Functional Status, Satisfaction With Overall Recovery, Satisfaction With Care Management, Care Transitions, and Caregiver Help (All LEJR Patients)

Domain	Measure	Range	CJR respondents (N)	Control respondents (N)	CJR risk-adjusted average	Control risk-adjusted average	Estimated difference	p-value
Change in functional status and pain[a]	Ability to walk by yourself without resting	-4 to 4	4,697	4,501	0.8	0.8	<0.1 (1.6%)	0.13
	Difficulty walking up or down 12 stairs	-3 to 3	4,530	4,327	0.8	0.8	<0.1 (0.5%)	0.66
	Difficulty rising from sitting	-4 to 4	4,782	4,575	1.2	1.2	<0.1 (0.1%)	0.91
	Difficulty standing	-4 to 4	4,788	4,579	1.2	1.2	<0.1 (0.5%)	0.39
	Use of a mobility aid	-2 to 2	4,762	4,556	0.2	0.1	<0.1 (0.9%)	0.34
	Difficulty getting on/off the toilet	-4 to 4	4,792	4,577	1.3	1.3	<0.1 (-0.4%)	0.47
	Frequency that pain interferes with normal activities	-4 to 4	4,778	4,576	2.0	2.0	<0.1 (-1.1%)	0.27
	Medication use for pain in the joint you had replaced	-3 to 3	4,698	4,491	0.6	0.6	<0.1 (0.0%)	0.96
Satisfaction with overall recovery[b]	Satisfaction with overall recovery since leaving the hospital	0 to 100	4,824	4,605	81.1	80.1	1.0	0.22
Satisfaction with care management[b]	Composite measure of satisfaction with care management	0 to 100	4,672	4,460	82.5	81.7	0.7	0.40
	Health care providers listened to preferences	0 to 100	4,817	4,607	78.6	77.2	1.4	0.17
	Satisfaction with discharge destination	0 to 100	4,779	4,579	83.3	82.2	1.1	0.18
	Satisfaction with care coordination	0 to 100	4,790	4,574	83.4	82.7	0.6	0.35
	Satisfaction with treatment instructions	0 to 100	4,831	4,628	83.9	83.7	0.2	0.73

Domain	Measure	Range	CJR respondents (N)	Control respondents (N)	CJR risk-adjusted average	Control risk-adjusted average	Estimated difference	p-value
Experience with care transitions[c]	Discharged from the hospital at the right time	0 to 100	4,776	4,555	87.5	88.0	-0.5	0.51
	Received the right amount of post-discharge care	0 to 100	4,815	4,605	85.2	85.4	-0.2	0.80
	Had all the medical equipment needed at home	0 to 100	4,578	4,366	91.6	92.0	-0.4	0.61
Caregiver help	Received any caregiver help[c]	0 to 100	4,555	4,350	96.0	96.6	-0.6	0.29
	Composite measure of caregiver help[d]	0 to 100	4,518	4,295	69.4	69.4	<.1	0.97
	Help needed putting on or taking off clothes[d]	0 to 100	4,568	4,357	60.8	61.4	-0.6	0.52
	Help needed bathing[d]	0 to 100	4,557	4,337	67.1	67.1	0.1	0.91
	Help needed using the toilet[d]	0 to 100	4,565	4,345	81.0	80.4	0.5	0.54

Source: CJR evaluation team analysis of patient survey data for episodes with inpatient discharge or outpatient procedure July, August, September, or October 2022.

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.

CJR = Comprehensive Care for Joint Replacement; LEJR = lower extremity joint replacement.

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

[b] Satisfaction outcomes are scaled from 0 to 100 points, where 0 = very dissatisfied, 25 = dissatisfied, 50 = neutral, 75 = satisfied, and 100 = very satisfied. The composite summarizes the level of satisfaction across the four measures of care management. Differences between CJR and control outcomes are reported in point terms.

[c] Indicates binary measure, reported as the percent of respondents reporting "Yes" to a given measure. Differences between CJR and control outcomes are reported in percentage point terms.

[d] Respondents were only asked about the amount of help needed with a given activity of daily living if they indicated that they received caregiver help. 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 outcomes are reported in point terms.

Exhibit E-6: Risk-Adjusted Survey-Based Results for Change in Functional Status, Satisfaction With Overall Recovery, Satisfaction With Care Management, Care Transitions, and Caregiver Help (Hip Fracture Patients)

Domain	Measure	Range	CJR respondents (N)	Control respondents (N)	CJR risk-adjusted average	Control risk-adjusted average	Estimated difference	p-value
Change in functional status and pain[a]	Ability to walk by yourself without resting	-4 to 4	512	582	-0.7	-0.8	0.1 (2.1%)	0.34
	Difficulty walking up or down 12 stairs	-3 to 3	475	548	-0.4	-0.6	0.2 (5.6%)	0.05
	Difficulty rising from sitting	-4 to 4	529	607	-0.3	-0.3	<0.1 (0.0%)	1.00
	Difficulty standing	-4 to 4	532	602	-0.3	-0.2	<0.1 (-1.2%)	0.46
	Use of a mobility aid	-2 to 2	523	603	-0.6	-0.7	0.1 (4.4%)	0.02
	Difficulty getting on/off the toilet	-4 to 4	527	606	0.0	-0.1	<0.1 (0.5%)	0.72
	Frequency that pain interferes with normal activities	-4 to 4	521	598	-0.3	-0.3	<0.1 (-0.7%)	0.67
	Medication use for pain in the joint you had replaced	-3 to 3	518	586	-0.2	-0.2	<0.1 (0.8%)	0.59
Satisfaction with overall recovery[b]	Satisfaction with overall recovery since leaving the hospital	0 to 100	577	642	75.8	71.5	4.3	0.08
Satisfaction with care management[b]	Composite measure of satisfaction with care management	0 to 100	543	608	77.4	73.1	4.3	0.03
	Health care providers listened to preferences	0 to 100	569	643	74.2	68.7	5.5	0.01
	Satisfaction with discharge destination	0 to 100	570	642	73.5	67.6	5.9	0.01
	Satisfaction with care coordination	0 to 100	560	624	76.2	73.3	2.9	0.35
	Satisfaction with treatment instructions	0 to 100	570	643	78.1	74.4	3.7	0.06
Experience with care transitions[c]	Discharged from the hospital at the right time	0 to 100	538	611	85.8	80.9	5.0	<0.01
	Received the right amount of post-discharge care	0 to 100	558	629	76.3	74.8	1.5	0.44
	Had all the medical equipment needed at home	0 to 100	525	577	86.9	87.2	-0.2	0.91

Domain	Measure	Range	CJR respondents (N)	Control respondents (N)	CJR risk-adjusted average	Control risk-adjusted average	Estimated difference	p-value
Caregiver help	Received any caregiver help[c]	0 to 100	519	573	95.5	95.6	<0.1	0.96
	Composite measure of caregiver help[d]	0 to 100	511	563	54.9	53.8	1.1	0.70
	Help needed putting on or taking off clothes[d]	0 to 100	522	575	49.6	49.4	0.1	0.95
	Help needed bathing[d]	0 to 100	522	571	48.2	48.6	0.4	0.84
	Help needed using the toilet[d]	0 to 100	522	570	64.7	64.2	0.5	0.88

Source: CJR evaluation team analysis of patient survey data for episodes with inpatient discharge in July, August, September, or October 2022.

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.

CJR = Comprehensive Care for Joint Replacement

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

[b] Satisfaction outcomes are scaled from 0 to 100 points, where 0 = very dissatisfied, 25 = dissatisfied, 50 = neutral, 75 = satisfied, and 100 = very satisfied. The composite summarizes the level of satisfaction across the four measures of care management. Differences between CJR and control outcomes are reported in point terms.

[c] Indicates binary measure, reported as the percent of respondents reporting "Yes" to a given measure. Differences between CJR and control outcomes are reported in percentage point terms.

[d] Respondents were only asked about the amount of help needed with a given activity of daily living if they indicated that they received caregiver help. 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 outcomes are reported in point terms.

Exhibit E-7: Risk-Adjusted Survey-Based Results for Change in Functional Status, Satisfaction With Overall Recovery, Satisfaction With Care Management, Care Transitions, and Caregiver Help (Elective Inpatient)

Domain	Measure	Range	CJR respondents (N)	Control respondents (N)	CJR risk-adjusted average	Control risk-adjusted average	Estimated difference	p-value
Change in functional status and pain[a]	Ability to walk by yourself without resting	-4 to 4	1,748	1,598	0.9	0.8	0.1 (3.2%)	0.14
	Difficulty walking up or down 12 stairs	-3 to 3	1,693	1,523	0.9	0.9	<0.1 (0.8%)	0.73
	Difficulty rising from sitting	-4 to 4	1,774	1,622	1.4	1.4	<0.1 (-0.4%)	0.75
	Difficulty standing	-4 to 4	1,777	1,624	1.3	1.2	<0.1 (1.3%)	0.27
	Use of a mobility aid	-2 to 2	1,771	1,615	0.2	0.2	<0.1 (2.2%)	0.16
	Difficulty getting on/off the toilet	-4 to 4	1,776	1,617	1.4	1.5	<0.1 (-1.5%)	0.29
	Frequency that pain interferes with normal activities	-4 to 4	1,779	1,623	2.2	2.3	<0.1 (-2.1%)	0.32
	Medication use for pain in the joint you had replaced	-3 to 3	1,748	1,598	0.7	0.6	0.1 (2.6%)	0.08
Satisfaction with overall recovery[b]	Satisfaction with overall recovery since leaving the hospital	0 to 100	1,780	1,627	80.0	78.2	1.8	0.17
Satisfaction with care management[b]	Composite measure of satisfaction with care management	0 to 100	1,730	1,568	82.4	80.7	1.7	0.14
	Health care providers listened to preferences	0 to 100	1,780	1,622	78.7	74.8	3.8	<0.01
	Satisfaction with discharge destination	0 to 100	1,769	1,610	82.9	81.6	1.3	0.41
	Satisfaction with care coordination	0 to 100	1,771	1,615	82.6	81.7	0.9	0.39
	Satisfaction with treatment instructions	0 to 100	1,787	1,634	83.9	83.1	0.8	0.38
Experience with care transitions[c]	Discharged from the hospital at the right time	0 to 100	1,766	1,604	88.5	89.5	-0.9	0.37
	Received the right amount of post-discharge care	0 to 100	1,778	1,633	86.2	86.1	0.1	0.91
	Had all the medical equipment needed at home	0 to 100	1,678	1,522	92.0	93.5	-1.6	0.15

Domain	Measure	Range	CJR respondents (N)	Control respondents (N)	CJR risk-adjusted average	Control risk-adjusted average	Estimated difference	p-value
Caregiver help	Received any caregiver help[c]	0 to 100	1,668	1,517	94.7	95.6	-0.9	0.33
	Composite measure of caregiver help[d]	0 to 100	1,656	1,502	70.6	72.4	-1.8	0.12
	Help needed putting on or taking off clothes[d]	0 to 100	1,669	1,524	62.9	64.5	-1.6	0.19
	Help needed bathing[d]	0 to 100	1,672	1,515	67.8	70.7	-2.9	0.09
	Help needed using the toilet[d]	0 to 100	1,670	1,519	81.6	82.5	-0.9	0.46

Source: CJR evaluation team analysis of patient survey data for episodes with inpatient discharge or outpatient procedure July, August, September, or October 2022.

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.

CJR = Comprehensive Care for Joint Replacement

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

[b] Satisfaction outcomes are scaled from 0 to 100 points, where 0 = very dissatisfied, 25 = dissatisfied, 50 = neutral, 75 = satisfied, and 100 = very satisfied. The composite summarizes the level of satisfaction across the four measures of care management. Differences between CJR and control outcomes are reported in point terms.

[c] Indicates binary measure, reported as the percent of respondents reporting "Yes" to a given measure. Differences between CJR and control outcomes are reported in percentage point terms.

[d] Respondents were only asked about the amount of help needed with a given activity of daily living if they indicated that they received caregiver help. 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 outcomes are reported in point terms.

Exhibit E-8: Risk-Adjusted Survey-Based Results for Change in Functional Status, Satisfaction With Overall Recovery, Satisfaction With Care Management, Care Transitions, and Caregiver Help (Elective Outpatient)

Domain	Measure	Range	CJR respondents (N)	Control respondents (N)	CJR risk-adjusted average	Control risk-adjusted average	Estimated difference	p-value
Change in functional status and pain[a]	Ability to walk by yourself without resting	-4 to 4	2,437	2,321	1.1	1.0	<0.1 (1.6%)	0.11
	Difficulty walking up or down 12 stairs	-3 to 3	2,362	2,256	1.0	1.0	<0.1 (-0.2%)	0.89
	Difficulty rising from sitting	-4 to 4	2,479	2,346	1.5	1.5	<0.1 (0.6%)	0.52
	Difficulty standing	-4 to 4	2,479	2,353	1.4	1.4	<0.1 (1.0%)	0.21
	Use of a mobility aid	-2 to 2	2,468	2,338	0.3	0.3	<0.1 (-0.3%)	0.70
	Difficulty getting on/off the toilet	-4 to 4	2,489	2,354	1.6	1.6	<0.1 (-0.2%)	0.74
	Frequency that pain interferes with normal activities	-4 to 4	2,478	2,355	2.4	2.4	<0.1 (-0.6%)	0.67
	Medication use for pain in the joint you had replaced	-3 to 3	2,432	2,307	0.7	0.7	<0.1 (-1.3%)	0.25
Satisfaction with overall recovery[b]	Satisfaction with overall recovery since leaving the hospital	0 to 100	2,467	2,336	82.5	82.4	0.1	0.94
Satisfaction with care management[b]	Composite measure of satisfaction with care management	0 to 100	2,399	2,284	83.8	83.8	<0.1	0.99
	Health care providers listened to preferences	0 to 100	2,468	2,342	79.2	79.8	-0.6	0.60
	Satisfaction with discharge destination	0 to 100	2,440	2,327	85.6	85.2	0.4	0.64
	Satisfaction with care coordination	0 to 100	2,459	2,335	84.9	84.9	<0.1	0.99
	Satisfaction with treatment instructions	0 to 100	2,474	2,351	85.2	85.3	-0.2	0.82
Experience with care transitions[c]	Discharged from the hospital at the right time	0 to 100	2,472	2,340	87.6	88.7	-1.1	0.18
	Received the right amount of post-discharge care	0 to 100	2,479	2,343	86.7	87.3	-0.6	0.48
	Had all the medical equipment needed at home	0 to 100	2,375	2,267	92.1	92.6	-0.5	0.64

Domain	Measure	Range	CJR respondents (N)	Control respondents (N)	CJR risk-adjusted average	Control risk-adjusted average	Estimated difference	p-value
Caregiver help	Received any caregiver help[c]	0 to 100	2,368	2,260	96.6	97.2	-0.6	0.40
	Composite measure of caregiver help[d]	0 to 100	2,351	2,230	72.0	71.4	0.6	0.60
	Help needed putting on or taking off clothes[d]	0 to 100	2,377	2,258	61.9	62.6	-0.7	0.59
	Help needed bathing[d]	0 to 100	2,363	2,251	70.8	69.0	1.8	0.29
	Help needed using the toilet[d]	0 to 100	2,373	2,256	83.8	82.8	1.1	0.33

Source: CJR evaluation team analysis of patient survey data for episodes with inpatient discharge or outpatient procedure July, August, September, or October 2022.

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.

CJR = Comprehensive Care for Joint Replacement

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

[b] Satisfaction outcomes are scaled from 0 to 100 points, where 0 = very dissatisfied, 25 = dissatisfied, 50 = neutral, 75 = satisfied, and 100 = very satisfied. The composite summarizes the level of satisfaction across the four measures of care management. Differences between CJR and control outcomes are reported in point terms.

[c] Indicates binary measure, reported as the percent of respondents reporting "Yes" to a given measure. Differences between CJR and control outcomes are reported in percentage point terms.

[d] Respondents were only asked about the amount of help needed with a given activity of daily living if they indicated that they received caregiver help. 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 outcomes are reported in point terms.

Exhibit E-9: Risk-Adjusted Survey-Based Results for Change in Functional Status, Satisfaction With Overall Recovery, Satisfaction With Care Management, Care Transitions, and Caregiver Help (Elective Inpatient Versus Outpatient)

Domain	Measure	Range	Elective inpatient model impact	p-value	Elective outpatient model impact	p-value	Differential impact of the CJR Model across settings	p-value
Change in functional status and pain[a]	Ability to walk by yourself without resting	-4 to 4	0.1 (3.2%)	0.14	<0.1 (1.6%)	0.11	<0.1 (-1.6%)	0.36
	Difficulty walking up or down 12 stairs	-3 to 3	<0.1 (0.8%)	0.73	<0.1 (-0.2%)	0.89	<0.1 (-0.9%)	0.62
	Difficulty rising from sitting	-4 to 4	<0.1 (-0.4%)	0.75	<0.1 (0.6%)	0.52	<0.1 (1.0%)	0.47
	Difficulty standing	-4 to 4	<0.1 (1.3%)	0.27	<0.1 (1.0%)	0.21	<0.1 (-0.3%)	0.82
	Use of a mobility aid	-2 to 2	<0.1 (2.2%)	0.16	<0.1 (-0.3%)	0.70	-0.1 (-2.5%)	0.08
	Difficulty getting on/off the toilet	-4 to 4	<0.1 (-1.5%)	0.29	<0.1 (-0.2%)	0.74	<0.1 (1.2%)	0.41
	Frequency that pain interferes with normal activities	-4 to 4	<0.1 (-2.1%)	0.32	<0.1 (-0.6%)	0.67	<0.1 (1.5%)	0.49
	Medication use for pain in the joint you had replaced	-3 to 3	0.1 (2.6%)	0.08	<0.1 (-1.3%)	0.25	-0.1 (-3.9%)	0.02
Satisfaction with overall recovery[b]	Satisfaction with overall recovery since leaving the hospital	0 to 100	1.8	0.17	0.1	0.94	-1.7	0.30
Satisfaction with care management[b]	Composite measure of satisfaction with care management	0 to 100	1.7	0.14	<0.1	0.99	-1.7	0.13
	Health care providers listened to preferences	0 to 100	3.8	<0.01	-0.6	0.60	-4.4	<0.01
	Satisfaction with discharge destination	0 to 100	1.3	0.41	0.4	0.64	-0.9	0.57
	Satisfaction with care coordination	0 to 100	0.9	0.39	<0.1	0.99	-0.9	0.46
	Satisfaction with treatment instructions	0 to 100	0.8	0.38	-0.2	0.82	-1.0	0.39
Experience with care transitions[c]	Discharged from the hospital at the right time	0 to 100	-0.9	0.37	-1.1	0.18	-0.2	0.89
	Received the right amount of post-discharge care	0 to 100	0.1	0.91	-0.6	0.48	-0.7	0.50
	Had all the medical equipment needed at home	0 to 100	-1.6	0.15	-0.5	0.64	1.1	0.44

Domain	Measure	Range	Elective inpatient model impact	p-value	Elective outpatient model impact	p-value	Differential impact of the CJR Model across settings	p-value
Caregiver help	Received any caregiver help[c]	0 to 100	-0.9	0.33	-0.6	0.40	0.3	0.77
	Composite measure of caregiver help[d]	0 to 100	-1.8	0.12	0.6	0.60	2.4	0.19
	Help needed putting on or taking off clothes[d]	0 to 100	-1.6	0.19	-0.7	0.59	0.9	0.66
	Help needed bathing[d]	0 to 100	-2.9	0.09	1.8	0.29	4.6	0.08
	Help needed using the toilet[d]	0 to 100	-0.9	0.46	1.1	0.33	2.0	0.29

Source: CJR evaluation team analysis of patient survey data for episodes with inpatient discharge or outpatient procedure July, August, September, or October 2022.

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.

CJR = Comprehensive Care for Joint Replacement

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

[b] Satisfaction outcomes are scaled from 0 to 100 points, where 0 = very dissatisfied, 25 = dissatisfied, 50 = neutral, 75 = satisfied, and 100 = very satisfied. The composite summarizes the level of satisfaction across the four measures of care management. Differences between CJR and control outcomes are reported in point terms.

[c] Indicates binary measure, reported as the percent of respondents reporting "Yes" to a given measure. Differences between CJR and control outcomes are reported in percentage point terms.

[d] Respondents were only asked about the amount of help needed with a given activity of daily living if they indicated that they received caregiver help. 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 outcomes are reported in point terms.

Exhibit E-10: Risk-Adjusted Survey-Based Results for Change in Functional Status, Satisfaction With Overall Recovery, Satisfaction With Care Management, Care Transitions, and Caregiver Help (All LEJR Patients, Excluding BPCI Advanced Episodes From the Control Group)

Domain	Measure	Range	CJR respondents (N)	Control respondents (N)	CJR risk-adjusted average	Control risk-adjusted average	Estimated difference	p-value
Change in functional status and pain[a]	Ability to walk by yourself without resting	-4 to 4	4,697	3,617	0.8	0.8	0.1 (2.6%)	0.02
	Difficulty walking up or down 12 stairs	-3 to 3	4,530	3,482	0.8	0.8	<0.1 (1.2%)	0.38
	Difficulty rising from sitting	-4 to 4	4,782	3,678	1.2	1.2	<0.1 (0.8%)	0.26
	Difficulty standing	-4 to 4	4,788	3,678	1.2	1.1	<0.1 (0.7%)	0.22
	Use of a mobility aid	-2 to 2	4,762	3,660	0.2	0.1	<0.1 (1.1%)	0.18
	Difficulty getting on/off the toilet	-4 to 4	4,792	3,673	1.3	1.3	<0.1 (0.0%)	0.96
	Frequency that pain interferes with normal activities	-4 to 4	4,778	3,684	2.0	2.0	<0.1 (-1.0%)	0.31
	Medication use for pain in the joint you had replaced	-3 to 3	4,698	3,624	0.6	0.6	<0.1 (-0.2%)	0.82
Satisfaction with overall recovery[b]	Satisfaction with overall recovery since leaving the hospital	0 to 100	4,824	3,710	81.0	79.9	1.1	0.19
Satisfaction with care management[b]	Composite measure of satisfaction with care management	0 to 100	4,672	3,585	82.3	81.8	0.5	0.58
	Health care providers listened to preferences	0 to 100	4,817	3,708	78.4	77.5	0.9	0.45
	Satisfaction with discharge destination	0 to 100	4,779	3,687	83.4	82.1	1.2	0.13
	Satisfaction with care coordination	0 to 100	4,790	3,679	83.3	82.2	1.1	0.20
	Satisfaction with treatment instructions	0 to 100	4,831	3,725	83.8	83.8	-0.1	0.93
Experience with care transitions[c]	Discharged from the hospital at the right time	0 to 100	4,776	3,661	87.5	88.0	-0.5	0.52
	Received the right amount of post-discharge care	0 to 100	4,815	3,703	85.3	85.6	-0.3	0.70
	Had all the medical equipment needed at home	0 to 100	4,578	3,512	91.5	92.7	-1.2	0.15

Domain	Measure	Range	CJR respondents (N)	Control respondents (N)	CJR risk-adjusted average	Control risk-adjusted average	Estimated difference	p-value
Caregiver help	Received any caregiver help[c]	0 to 100	4,555	3,501	95.8	96.6	-0.8	0.19
	Composite measure of caregiver help[d]	0 to 100	4,518	3,450	69.4	69.1	0.2	0.82
	Help needed putting on or taking off clothes[d]	0 to 100	4,568	3,504	60.8	61.2	-0.3	0.78
	Help needed bathing[d]	0 to 100	4,557	3,487	67.1	66.7	0.4	0.74
	Help needed using the toilet[d]	0 to 100	4,565	3,491	81.0	80.4	0.6	0.55

Source: CJR evaluation team analysis of patient survey data for episodes with inpatient discharge or outpatient procedure July, August, September, or October 2022.

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.

CJR = Comprehensive Care for Joint Replacement; LEJR = lower extremity joint replacement.

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

[b] Satisfaction outcomes are scaled from 0 to 100 points, where 0 = very dissatisfied, 25 = dissatisfied, 50 = neutral, 75 = satisfied, and 100 = very satisfied. The composite summarizes the level of satisfaction across the four measures of care management. Differences between CJR and control outcomes are reported in point terms.

[c] Indicates binary measure, reported as the percent of respondents reporting "Yes" to a given measure. Differences between CJR and control outcomes are reported in percentage point terms.

[d] Respondents were only asked about the amount of help needed with a given activity of daily living if they indicated that they received caregiver help. 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 outcomes are reported in point terms.

Exhibit E-11: Risk-Adjusted Survey-Based Results for Change in Functional Status, Satisfaction With Overall Recovery, Satisfaction With Care Management, Care Transitions, and Caregiver Help (Hip Fracture Patients, Excluding BPCI Advanced Episodes From the Control Group)

Domain	Measure	Range	CJR respondents (N)	Control respondents (N)	CJR risk-adjusted average	Control risk-adjusted average	Estimated difference	p-value
Change in functional status and pain[a]	Ability to walk by yourself without resting	-4 to 4	512	505	-0.7	-0.8	0.1 (4.1%)	0.10
	Difficulty walking up or down 12 stairs	-3 to 3	475	477	-0.4	-0.6	0.2 (6.5%)	0.02
	Difficulty rising from sitting	-4 to 4	529	528	-0.3	-0.3	<0.1 (0.6%)	0.70
	Difficulty standing	-4 to 4	532	524	-0.3	-0.2	<0.1 (-0.4%)	0.84
	Use of a mobility aid	-2 to 2	523	525	-0.6	-0.7	0.1 (2.7%)	0.15
	Difficulty getting on/off the toilet	-4 to 4	527	526	<0.1	-0.1	0.1 (1.3%)	0.47
	Frequency that pain interferes with normal activities	-4 to 4	521	524	-0.3	-0.3	<0.1 (0.0%)	0.98
	Medication use for pain in the joint you had replaced	-3 to 3	518	511	-0.2	-0.2	<0.1 (0.9%)	0.60
Satisfaction with overall recovery[b]	Satisfaction with overall recovery since leaving the hospital	0 to 100	577	559	75.4	71.2	4.2	0.10
Satisfaction with care management[b]	Composite measure of satisfaction with care management	0 to 100	543	529	77.5	73.0	4.5	0.03
	Health care providers listened to preferences	0 to 100	569	561	74.0	67.7	6.3	0.01
	Satisfaction with discharge destination	0 to 100	570	559	73.5	68.1	5.4	0.02
	Satisfaction with care coordination	0 to 100	560	542	75.9	72.9	3.0	0.36
	Satisfaction with treatment instructions	0 to 100	570	560	77.8	73.5	4.3	0.07
Experience with care transitions[c]	Discharged from the hospital at the right time	0 to 100	538	533	85.7	80.8	4.9	0.01
	Received the right amount of post-discharge care	0 to 100	558	549	76.5	74.9	1.5	0.37
	Had all the medical equipment needed at home	0 to 100	525	503	86.8	87.4	-0.6	0.79

Domain	Measure	Range	CJR respondents (N)	Control respondents (N)	CJR risk-adjusted average	Control risk-adjusted average	Estimated difference	p-value
Caregiver help	Received any caregiver help[c]	0 to 100	519	498	95.5	96.0	-0.5	0.69
	Composite measure of caregiver help[d]	0 to 100	511	490	54.8	53.1	1.6	0.57
	Help needed putting on or taking off clothes[d]	0 to 100	522	500	49.6	48.7	1.0	0.71
	Help needed bathing[d]	0 to 100	522	496	48.6	47.2	1.4	0.54
	Help needed using the toilet[d]	0 to 100	522	495	65.1	63.2	1.9	0.56

Source: CJR evaluation team analysis of patient survey data for episodes with inpatient discharge in July, August, September, or October 2022.

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.

CJR = Comprehensive Care for Joint Replacement

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

[b] Satisfaction outcomes are scaled from 0 to 100 points, where 0 = very dissatisfied, 25 = dissatisfied, 50 = neutral, 75 = satisfied, and 100 = very satisfied. The composite summarizes the level of satisfaction across the four measures of care management. Differences between CJR and control outcomes are reported in point terms.

[c] Indicates binary measure, reported as the percent of respondents reporting "Yes" to a given measure. Differences between CJR and control outcomes are reported in percentage point terms.

[d] Respondents were only asked about the amount of help needed with a given activity of daily living if they indicated that they received caregiver help. 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 outcomes are reported in point terms.

Exhibit E-12: Risk-Adjusted Survey-Based Results for Change in Functional Status, Satisfaction With Overall Recovery, Satisfaction With Care Management, Care Transitions, and Caregiver Help (Elective Inpatient, Excluding BPCI Advanced Episodes From the Control Group)

Domain	Measure	Range	CJR respondents (N)	Control respondents (N)	CJR risk-adjusted average	Control risk-adjusted average	Estimated difference	p-value
Change in functional status and pain[a]	Ability to walk by yourself without resting	-4 to 4	1,748	1,251	0.9	0.8	0.1 (4.9%)	0.01
	Difficulty walking up or down 12 stairs	-3 to 3	1,693	1,188	0.9	0.8	0.1 (2.6%)	0.23
	Difficulty rising from sitting	-4 to 4	1,774	1,274	1.4	1.3	<0.1 (1.1%)	0.31
	Difficulty standing	-4 to 4	1,777	1,272	1.3	1.2	0.1 (2.6%)	0.01
	Use of a mobility aid	-2 to 2	1,771	1,264	0.2	0.2	0.1 (2.8%)	0.08
	Difficulty getting on/off the toilet	-4 to 4	1,776	1,264	1.4	1.5	<0.1 (-0.7%)	0.56
	Frequency that pain interferes with normal activities	-4 to 4	1,779	1,273	2.2	2.2	<0.1 (0.4%)	0.83
	Medication use for pain in the joint you had replaced	-3 to 3	1,748	1,256	0.7	0.6	0.1 (3.4%)	0.03
Satisfaction with overall recovery[b]	Satisfaction with overall recovery since leaving the hospital	0 to 100	1,780	1,277	79.8	77.9	1.9	0.21
Satisfaction with care management[b]	Composite measure of satisfaction with care management	0 to 100	1,730	1,230	81.9	81.0	0.9	0.45
	Health care providers listened to preferences	0 to 100	1,780	1,274	78.0	75.5	2.6	0.13
	Satisfaction with discharge destination	0 to 100	1,769	1,261	82.4	81.9	0.5	0.80
	Satisfaction with care coordination	0 to 100	1,771	1,267	82.4	81.3	1.1	0.41
	Satisfaction with treatment instructions	0 to 100	1,787	1,280	83.5	83.3	0.2	0.87
Experience with care transitions[c]	Discharged from the hospital at the right time	0 to 100	1,766	1,253	88.6	89.8	-1.2	0.31
	Received the right amount of post-discharge care	0 to 100	1,778	1,277	86.1	85.8	0.4	0.77
	Had all the medical equipment needed at home	0 to 100	1,678	1,196	92.0	93.0	-1.0	0.40

Domain	Measure	Range	CJR respondents (N)	Control respondents (N)	CJR risk-adjusted average	Control risk-adjusted average	Estimated difference	p-value
Caregiver help	Received any caregiver help[c]	0 to 100	1,668	1,193	94.7	95.7	-1.0	0.21
	Composite measure of caregiver help[d]	0 to 100	1,656	1,174	70.7	72.3	-1.5	0.22
	Help needed putting on or taking off clothes[d]	0 to 100	1,669	1,195	62.9	64.1	-1.2	0.37
	Help needed bathing[d]	0 to 100	1,672	1,186	67.6	71.2	-3.6	0.05
	Help needed using the toilet[d]	0 to 100	1,670	1,190	81.6	82.3	-0.6	0.60

Source: CJR evaluation team analysis of patient survey data for episodes with inpatient discharge or outpatient procedure July, August, September, or October 2022.

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.

CJR = Comprehensive Care for Joint Replacement.

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

[b] Satisfaction outcomes are scaled from 0 to 100 points, where 0 = very dissatisfied, 25 = dissatisfied, 50 = neutral, 75 = satisfied, and 100 = very satisfied. The composite summarizes the level of satisfaction across the four measures of care management. Differences between CJR and control outcomes are reported in point terms.

[c] Indicates binary measure, reported as the percent of respondents reporting "Yes" to a given measure. Differences between CJR and control outcomes are reported in percentage point terms.

[d] Respondents were only asked about the amount of help needed with a given activity of daily living if they indicated that they received caregiver help. 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 outcomes are reported in point terms.

Exhibit E-13: Risk-Adjusted Survey-Based Results for Change in Functional Status, Satisfaction With Overall Recovery, Satisfaction With Care Management, Care Transitions, and Caregiver Help (Elective Outpatient, Excluding BPCI Advanced Episodes From the Control Group)

Domain	Measure	Range	CJR respondents (N)	Control respondents (N)	CJR risk-adjusted average	Control risk-adjusted average	Estimated difference	p-value
Change in functional status and pain[a]	Ability to walk by yourself without resting	-4 to 4	2,437	1,861	1.1	1.0	0.1 (2.0%)	0.06
	Difficulty walking up or down 12 stairs	-3 to 3	2,362	1,817	1.0	1.0	<0.1 (0.0%)	0.99
	Difficulty rising from sitting	-4 to 4	2,479	1,876	1.5	1.5	<0.1 (0.7%)	0.47
	Difficulty standing	-4 to 4	2,479	1,882	1.4	1.4	<0.1 (0.6%)	0.34
	Use of a mobility aid	-2 to 2	2,468	1,871	0.3	0.3	<0.1 (-0.1%)	0.94
	Difficulty getting on/off the toilet	-4 to 4	2,489	1,883	1.5	1.6	<0.1 (-0.2%)	0.79
	Frequency that pain interferes with normal activities	-4 to 4	2,478	1,887	2.3	2.4	<0.1 (-1.4%)	0.37
	Medication use for pain in the joint you had replaced	-3 to 3	2,432	1,857	0.7	0.7	<0.1 (-1.7%)	0.09
Satisfaction with overall recovery[b]	Satisfaction with overall recovery since leaving the hospital	0 to 100	2,467	1,874	82.5	82.6	<0.1	0.98
Satisfaction with care management[b]	Composite measure of satisfaction with care management	0 to 100	2,399	1,826	83.7	84.4	-0.7	0.53
	Health care providers listened to preferences	0 to 100	2,468	1,873	79.2	80.4	-1.2	0.35
	Satisfaction with discharge destination	0 to 100	2,440	1,867	85.4	85.7	-0.3	0.77
	Satisfaction with care coordination	0 to 100	2,459	1,870	85.0	84.4	0.6	0.57
	Satisfaction with treatment instructions	0 to 100	2,474	1,885	85.1	85.9	-0.8	0.37
Experience with care transitions[c]	Discharged from the hospital at the right time	0 to 100	2,472	1,875	87.5	88.7	-1.2	0.15
	Received the right amount of post-discharge care	0 to 100	2,479	1,877	86.8	87.5	-0.8	0.35
	Had all the medical equipment needed at home	0 to 100	2,375	1,813	92.1	93.6	-1.5	0.10

Domain	Measure	Range	CJR respondents (N)	Control respondents (N)	CJR risk-adjusted average	Control risk-adjusted average	Estimated difference	p-value
Caregiver help	Received any caregiver help[c]	0 to 100	2,368	1,810	96.3	97.0	-0.7	0.31
	Composite measure of caregiver help[d]	0 to 100	2,351	1,786	72.0	71.1	0.9	0.48
	Help needed putting on or taking off clothes[d]	0 to 100	2,377	1,809	61.9	62.5	-0.6	0.64
	Help needed bathing[d]	0 to 100	2,363	1,805	70.6	68.5	2.1	0.20
	Help needed using the toilet[d]	0 to 100	2,373	1,806	83.9	82.7	1.2	0.31

Source: CJR evaluation team analysis of patient survey data for episodes with inpatient discharge or outpatient procedure July, August, September, or October 2022.

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.

CJR = Comprehensive Care for Joint Replacement

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

[b] Satisfaction outcomes are scaled from 0 to 100 points, where 0 = very dissatisfied, 25 = dissatisfied, 50 = neutral, 75 = satisfied, and 100 = very satisfied. The composite summarizes the level of satisfaction across the four measures of care management. Differences between CJR and control outcomes are reported in point terms.

[c] Indicates binary measure, reported as the percent of respondents reporting "Yes" to a given measure. Differences between CJR and control outcomes are reported in percentage point terms.

[d] Respondents were only asked about the amount of help needed with a given activity of daily living if they indicated that they received caregiver help. 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 outcomes are reported in point terms.

Exhibit E-14: Risk-Adjusted Survey-Based Results for Change in Functional Status, Satisfaction With Overall Recovery, Satisfaction With Care Management, Care Transitions, and Caregiver Help (Elective Inpatient Versus Outpatient, Excluding BPCI Advanced Episodes From the Control Group)

Domain	Measure	Range	Elective inpatient model impact	p-value	Elective outpatient model impact	p-value	Differential impact of the CJR Model across settings	p-value
Change in functional status and pain[a]	Ability to walk by yourself without resting	-4 to 4	0.1 (4.9%)	0.01	0.1 (2.0%)	0.06	-0.1 (-2.9%)	0.09
	Difficulty walking up or down 12 stairs	-3 to 3	0.1 (2.6%)	0.23	<0.1 (0.0%)	0.99	-0.1 (-2.6%)	0.12
	Difficulty rising from sitting	-4 to 4	<0.1 (1.1%)	0.31	<0.1 (0.7%)	0.47	<0.1 (-0.4%)	0.77
	Difficulty standing	-4 to 4	0.1 (2.6%)	<0.01	<0.1 (0.6%)	0.34	-0.1 (-1.9%)	0.04
	Use of a mobility aid	-2 to 2	0.1 (2.8%)	0.08	<0.1 (-0.1%)	0.94	-0.1 (-2.8%)	0.06
	Difficulty getting on/off the toilet	-4 to 4	<0.1 (-0.7%)	0.56	<0.1 (-0.2%)	0.79	<0.1 (0.5%)	0.70
	Frequency that pain interferes with normal activities	-4 to 4	<0.1 (0.4%)	0.83	<0.1 (-1.4%)	0.37	<0.1 (-1.7%)	0.47
	Medication use for pain in the joint you had replaced	-3 to 3	0.1 (3.4%)	0.03	<0.1 (-1.7%)	0.09	-0.1 (-5.1%)	<0.01
Satisfaction with overall recovery[b]	Satisfaction with overall recovery since leaving the hospital	0 to 100	1.9	0.21	<0.1	0.98	-1.9	0.29
Satisfaction with care management[b]	Composite measure of satisfaction with care management	0 to 100	0.9	0.45	-0.7	0.53	-1.6	0.20
	Health care providers listened to preferences	0 to 100	2.6	0.13	-1.2	0.35	-3.8	0.02
	Satisfaction with discharge destination	0 to 100	0.5	0.80	-0.3	0.77	-0.8	0.66
	Satisfaction with care coordination	0 to 100	1.1	0.41	0.6	0.57	-0.4	0.72
	Satisfaction with treatment instructions	0 to 100	0.2	0.87	-0.8	0.37	-1.0	0.45
Experience with care transitions[c]	Discharged from the hospital at the right time	0 to 100	-1.2	0.31	-1.2	0.15	-0.1	0.97
	Received the right amount of post-discharge care	0 to 100	0.4	0.77	-0.8	0.35	-1.1	0.31
	Had all the medical equipment needed at home	0 to 100	-1.0	0.40	-1.5	0.10	-0.6	0.67

Domain	Measure	Range	Elective inpatient model impact	p-value	Elective outpatient model impact	p-value	Differential impact of the CJR Model across settings	p-value
Caregiver help	Received any caregiver help[c]	0 to 100	-1.0	0.21	-0.7	0.31	0.3	0.70
	Composite measure of caregiver help[d]	0 to 100	-1.5	0.22	0.9	0.48	2.4	0.23
	Help needed putting on or taking off clothes[d]	0 to 100	-1.2	0.37	-0.6	0.64	0.5	0.81
	Help needed bathing[d]	0 to 100	-3.6	0.05	2.1	0.20	5.7	0.04
	Help needed using the toilet[d]	0 to 100	-0.6	0.60	1.2	0.31	1.8	0.36

Source: CJR evaluation team analysis of patient survey data for episodes with inpatient discharge or outpatient procedure July, August, September, or October 2022.

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.

CJR = Comprehensive Care for Joint Replacement

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

[b] Satisfaction outcomes are scaled from 0 to 100 points, where 0 = very dissatisfied, 25 = dissatisfied, 50 = neutral, 75 = satisfied, and 100 = very satisfied. The composite summarizes the level of satisfaction across the four measures of care management. Differences between CJR and control outcomes are reported in point terms.

[c] Indicates binary measure, reported as the percent of respondents reporting "Yes" to a given measure. Differences between CJR and control outcomes are reported in percentage point terms.

[d] Respondents were only asked about the amount of help needed with a given activity of daily living if they indicated that they received caregiver help. 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 outcomes are reported in point terms.

Exhibit E-15: Unweighted and Unadjusted Correlation Coefficients Between Outcome Measures and Demographic Variables (All LEJR Patients)

Domain	Measure	Age	Female sex	Black or African Am.	Hisp.	Non-Hisp. White	Other race	Missing race/ ethn.	Some high school	High school or GED	College – less than BA	BA degree	More than BA degree	Missing educ.
Change in functional status and pain[a]	Ability to walk by yourself without resting	-0.232	-0.011	-0.035	0.016	0.020	-0.002	-0.012	-0.033	-0.049	0.026	0.023	0.031	-0.037
	Difficulty walking up or down 12 stairs	-0.234	-0.007	-0.019	0.002	0.016	-0.014	-0.003	-0.019	-0.045	0.038	0.023	0.013	-0.046
	Difficulty rising from sitting	-0.263	0.019	-0.008	0.025	-0.005	-0.015	0.005	-0.025	-0.010	0.040	-0.005	-0.015	-0.011
	Difficulty standing	-0.239	0.040	-0.005	0.020	0.005	-0.014	-0.008	-0.019	-0.010	0.044	0.001	-0.027	-0.010
	Use of a mobility aid	-0.159	0.009	-0.023	-0.023	0.025	-0.012	0.002	-0.025	-0.045	0.041	0.002	0.032	-0.047
	Difficulty getting on/off the toilet	-0.216	0.051	0.005	0.024	-0.008	-0.018	0.003	-0.013	-0.001	0.039	-0.002	-0.031	-0.008
	Frequency that pain interferes with normal activities	-0.218	-0.024	-0.023	0.007	0.019	-0.025	-0.001	-0.024	-0.024	0.036	0.013	0.003	-0.041
	Medication use for pain in the joint you had replaced	-0.123	0.025	-0.027	-0.024	0.042	-0.030	-0.008	-0.029	-0.025	0.013	0.028	0.008	-0.022
Satisfaction with overall recovery[b]	Satisfaction with overall recovery since leaving the hospital	-0.065	0.026	-0.011	-0.014	0.050	-0.011	-0.045	-0.015	-0.003	-0.017	0.009	0.038	-0.036
Satisfaction with care manage-ment[b]	Composite measure of satisfaction with care management	0.120	0.014	0.022	0.012	-0.056	0.003	0.051	0.014	0.023	0.008	-0.022	-0.041	0.045
	Health care providers listened to preferences	-0.100	-0.013	-0.012	-0.002	0.036	0.005	-0.040	-0.006	-0.031	-0.001	0.027	0.029	-0.036
	Satisfaction with discharge destination	-0.125	-0.017	-0.033	-0.019	0.056	-0.008	-0.037	-0.027	-0.032	-0.009	0.024	0.052	-0.040
	Satisfaction with care coordination	-0.098	-0.018	-0.014	-0.021	0.054	-0.005	-0.046	-0.010	-0.008	-0.002	0.007	0.026	-0.035
	Satisfaction with treatment instructions	-0.096	-0.009	-0.008	-0.009	0.052	-0.004	-0.056	-0.004	-0.013	-0.004	0.017	0.025	-0.039

Domain	Measure	Age	Female sex	Black or African Am.	Hisp.	Non-Hisp. White	Other race	Missing race/ ethn.	Some high school	High school or GED	College – less than BA	BA degree	More than BA degree	Missing educ.
Experience with care transitions[c]	Discharged from the hospital at the right time	-0.027	-0.077	-0.048	-0.026	0.082	-0.033	-0.045	-0.023	-0.002	-0.008	0.021	0.027	-0.052
	Received the right amount of post-discharge care	-0.043	-0.033	-0.029	0.009	0.030	0.012	-0.031	-0.012	0.005	-0.007	0.013	0.002	-0.012
	Had all the medical equipment needed at home	0.020	0.011	-0.028	-0.040	0.077	-0.023	-0.049	-0.013	0.025	-0.016	0.028	-0.023	-0.009
Caregiver help	Received any caregiver help[c]	-0.022	-0.031	-0.002	-0.005	0.026	-0.010	-0.025	0.016	0.008	-0.004	0.003	-0.002	-0.021
	Composite measure of caregiver help[d]	-0.043	-0.016	-0.044	-0.082	0.087	-0.049	-0.013	-0.050	-0.033	0.005	0.033	0.040	-0.043
	Help needed putting on or taking off clothes[d]	-0.015	0.075	-0.023	-0.062	0.051	-0.023	-0.004	-0.033	-0.011	-0.003	0.020	0.022	-0.025
	Help needed bathing[d]	-0.058	-0.057	-0.044	-0.065	0.066	-0.039	-0.001	-0.041	-0.034	0.009	0.030	0.032	-0.034
	Help needed using the toilet[d]	-0.038	-0.056	-0.047	-0.081	0.104	-0.062	-0.028	-0.054	-0.036	0.006	0.034	0.044	-0.047

Source: CJR evaluation team analysis of patient survey data for episodes with inpatient discharge or outpatient procedure July, August, September, or October 2022.

Notes: The coefficients in this exhibit are raw correlations (unweighted and unadjusted) between variables in the cross-sectional all-LEJR sample. Other race indicates a response of "Asian," "Native Hawaiian or Other Pacific Islander," or "American Indian or Alaska Native."

CJR = Comprehensive Care for Joint Replacement; GED = General Educational Development; LEJR = lower extremity joint replacement; BA = bachelor's.

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

[b] Satisfaction outcomes are scaled from 0 to 100 points, where 0 = very dissatisfied, 25 = dissatisfied, 50 = neutral, 75 = satisfied, and 100 = very satisfied. The composite summarizes the level of satisfaction across the four measures of care management. Differences between CJR and control outcomes are reported in point terms.

[c] Indicates binary measure, reported as the percent of respondents reporting "Yes" to a given measure. Differences between CJR and control outcomes are reported in percentage point terms.

[d] Respondents were only asked about the amount of help needed with a given activity of daily living if they indicated that they received caregiver help. 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 outcomes are reported in point terms.

Exhibit E-16: Unweighted and Unadjusted Correlation Coefficients Between Outcome Measures and Procedure Characteristics, Medicare Enrollment Status, Prior Health Conditions, and Proxy Status (All LEJR Patients)

Domain	Measure	Fracture	Knee	MS-DRG 469 or 521	HCC score	SNF/NH stay	Dual eligibility	Disabled	ACO assignment	Pre-hospital functional status	Proxy status	Missing proxy status
Change in functional status and pain[a]	Ability to walk by yourself without resting	-0.397	0.074	-0.127	-0.161	-0.048	-0.033	-0.007	0.006	-0.499	-0.087	-0.051
	Difficulty walking up or down 12 stairs	-0.444	0.130	-0.170	-0.183	-0.050	-0.041	-0.002	-0.010	-0.484	-0.122	-0.041
	Difficulty rising from sitting	-0.447	0.085	-0.152	-0.166	-0.031	0.002	0.043	0.000	-0.597	-0.107	-0.038
	Difficulty standing	-0.414	0.083	-0.139	-0.166	-0.042	-0.002	0.025	0.006	-0.587	-0.117	-0.035
	Use of a mobility aid	-0.323	0.042	-0.123	-0.077	-0.021	-0.010	0.016	0.025	-0.445	-0.079	-0.052
	Difficulty getting on/off the toilet	-0.417	0.075	-0.132	-0.124	-0.033	0.006	0.053	0.007	-0.656	-0.107	-0.035
	Frequency that pain interferes with normal activities	-0.546	0.091	-0.179	-0.169	-0.037	-0.018	-0.004	0.003	-0.573	-0.132	-0.049
	Medication use for pain in the joint you had replaced	-0.243	-0.012	-0.066	-0.077	-0.014	-0.014	0.028	0.015	-0.340	-0.050	-0.039
Satisfaction with overall recovery[b]	Satisfaction with overall recovery since leaving the hospital	-0.052	-0.014	-0.071	-0.092	-0.041	-0.029	-0.038	0.002	-0.008	-0.045	-0.050
Satisfaction with care manage-ment[b]	Composite measure of satisfaction with care management	0.088	-0.025	0.086	0.107	0.031	0.050	0.026	-0.011	0.000	0.060	0.061
	Health care providers listened to preferences	-0.052	0.011	-0.058	-0.087	-0.023	-0.036	-0.019	0.001	0.004	-0.040	-0.058
	Satisfaction with discharge destination	-0.099	0.039	-0.088	-0.111	-0.047	-0.057	-0.026	0.007	0.017	-0.063	-0.060
	Satisfaction with care coordination	-0.075	0.015	-0.079	-0.083	-0.010	-0.042	-0.016	0.004	-0.015	-0.053	-0.038
	Satisfaction with treatment instructions	-0.078	0.024	-0.079	-0.080	-0.024	-0.035	-0.017	0.019	-0.015	-0.054	-0.047

Domain	Measure	Fracture	Knee	MS-DRG 469 or 521	HCC score	SNF/NH stay	Dual eligibility	Disabled	ACO assignment	Pre-hospital functional status	Proxy status	Missing proxy status
Experience with care transitions[c]	Discharged from the hospital at the right time	-0.007	-0.007	-0.007	-0.066	-0.025	-0.050	-0.074	-0.003	0.093	-0.057	-0.025
	Received the right amount of post-discharge care	-0.056	0.000	-0.072	-0.100	-0.056	-0.018	-0.052	-0.001	0.044	-0.050	-0.006
	Had all the medical equipment needed at home	-0.047	-0.006	-0.044	-0.036	0.001	-0.048	-0.032	0.007	0.013	-0.019	-0.001
Caregiver help	Received any caregiver help[c]	-0.019	0.023	-0.001	-0.038	-0.042	-0.042	-0.022	0.016	0.001	0.025	-0.019
	Composite measure of caregiver help[d]	-0.170	0.089	-0.115	-0.156	-0.071	-0.070	-0.096	0.012	0.151	-0.203	-0.021
	Help needed putting on or taking off clothes[d]	-0.110	0.124	-0.070	-0.113	-0.032	-0.037	-0.068	0.004	0.123	-0.137	-0.017
	Help needed bathing[d]	-0.172	0.067	-0.109	-0.151	-0.069	-0.055	-0.089	0.004	0.121	-0.172	-0.009
	Help needed using the toilet[d]	-0.147	0.035	-0.115	-0.137	-0.081	-0.088	-0.088	0.021	0.142	-0.206	-0.028

Source: CJR evaluation team analysis of patient survey data for episodes with inpatient discharge or outpatient procedure July, August, September, or October 2022.

Notes: The coefficients in this exhibit are raw correlations (unweighted and unadjusted) between variables in the cross-sectional all-LEJR sample.

CJR = Comprehensive Care for Joint Replacement; LEJR = lower extremity joint replacement; MS-DRG = Medicare Severity-Diagnosis Related Group; HCC = hierarchical condition category; SNF/NH = skilled nursing facility or nursing home; Disabled = Originally qualified for Medicare due to disability; ACO = Accountable Care Organization; Proxy status = patient had help from someone else in responding to the survey.

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

[b] Satisfaction outcomes are scaled from 0 to 100 points, where 0 = very dissatisfied, 25 = dissatisfied, 50 = neutral, 75 = satisfied, and 100 = very satisfied. The composite summarizes the level of satisfaction across the four measures of care management. Differences between CJR and control outcomes are reported in point terms.

[c] Indicates binary measure, reported as the percent of respondents reporting "Yes" to a given measure. Differences between CJR and control outcomes are reported in percentage point terms.

[d] Respondents were only asked about the amount of help needed with a given activity of daily living if they indicated that they received caregiver help. 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 outcomes are reported in point terms.

D. Health Equity Analyses – Results Tables

1. Functional Status

Exhibit E-17: Risk-Adjusted Survey-Based Results for Differences in Functional Status Between Patients Who Were Dually Eligible and Patients Who Were Not Dually Eligible Discharged From Mandatory Hospitals

Survey question:	Survey response	Dually eligible				Non-dually eligible				Differential effect	p-value
		CJR risk-adjusted rates (%)	Control risk-adjusted rates (%)	Difference (pp)	p-value	CJR risk-adjusted rates (%)	Control risk-adjusted rates (%)	Difference (pp)	p-value		
How much does pain in the joint interfere with your normal activities?	All of the time	4.5%	3.6%	0.9	0.33	1.7%	1.6%	0.1	0.42	0.9	0.36
	Most of the time	10.2%	8.4%	1.8	0.32	4.4%	4.2%	0.1	0.42	1.7	0.36
	Some of the time	28.4%	25.6%	2.7	0.33	16.9%	16.4%	0.4	0.42	2.3	0.41
	A little of the time	34.1%	35.2%	-1.2	0.33	34.8%	34.6%	0.2	0.42	-1.4	0.25
	None of the time	22.8%	27.1%	-4.3	0.33	42.2%	43.1%	-0.9	0.42	-3.5	0.45
Are you taking any of the following types of medications for the joint you had replaced?	Prescription and OTC	19.9%	14.9%	5.0	0.13	8.4%	9.0%	-0.6	0.15	5.6	0.10
	Prescription only	11.3%	9.4%	1.8	0.14	6.3%	6.6%	-0.3	0.15	2.2	0.08
	OTC only	30.5%	29.3%	1.2	0.23	24.6%	25.3%	-0.7	0.15	1.9	0.09
	None	38.3%	46.4%	-8.1	0.14	60.8%	59.2%	1.6	0.15	-9.7	0.08
What best describes your use of a mobility aid such as a wheelchair, scooter, walker, or cane?	Always Used	39.4%	38.5%	1.0	0.81	13.2%	13.2%	0.0	0.98	1.0	0.80
	Sometimes Used	30.9%	31.0%	-0.1	0.81	22.6%	22.6%	0.0	0.98	-0.1	0.90
	Never Used	29.7%	30.6%	-0.9	0.80	64.2%	64.2%	0.0	0.98	-0.9	0.81
What best describes your ability to walk by yourself without resting?	Unable to walk	19.3%	17.3%	2.0	0.42	5.1%	5.4%	-0.3	0.07	2.4	0.35
	Walk from one room to another	32.7%	31.3%	1.4	0.43	15.2%	15.8%	-0.7	0.07	2.0	0.25
	Walk one block	21.7%	22.4%	-0.7	0.42	19.0%	19.4%	-0.4	0.07	-0.2	0.77
	Walk several blocks	15.1%	16.3%	-1.2	0.42	23.8%	23.8%	0.0	0.12	-1.3	0.41
	Walk more than several blocks	11.2%	12.7%	-1.5	0.43	37.0%	35.6%	1.4	0.07	-2.9	0.16

Survey question:	Survey response	Dually eligible				Non-dually eligible				Differential effect	p-value
		CJR risk-adjusted rates (%)	Control risk-adjusted rates (%)	Difference (pp)	p-value	CJR risk-adjusted rates (%)	Control risk-adjusted rates (%)	Difference (pp)	p-value		
How much difficulty do you have walking up or down 12 stairs?	I am not able to	25.3%	24.9%	0.4	0.92	8.4%	8.6%	-0.1	0.64	0.5	0.89
	A lot of difficulty	15.7%	15.6%	0.1	0.92	8.3%	8.4%	-0.1	0.64	0.2	0.84
	Some difficulty	43.7%	43.9%	-0.2	0.92	46.6%	46.8%	-0.2	0.64	0.0	0.99
	No difficulty	15.3%	15.6%	-0.3	0.92	36.7%	36.2%	0.4	0.64	-0.7	0.81
How much difficulty do you have rising from sitting?	Extreme	3.7%	2.8%	0.9	0.23	1.0%	1.1%	-0.1	0.25	0.9	0.20
	Severe	6.9%	5.5%	1.4	0.22	2.2%	2.3%	-0.1	0.26	1.5	0.18
	Moderate	37.5%	33.4%	4.1	0.22	19.2%	19.9%	-0.7	0.25	4.8	0.16
	Mild	34.2%	36.4%	-2.2	0.21	36.6%	36.9%	-0.3	0.25	-1.9	0.28
	None	17.6%	21.8%	-4.3	0.22	40.9%	39.8%	1.1	0.25	-5.4	0.14
How much difficulty do you have standing?	Extreme	3.9%	3.3%	0.5	0.45	0.8%	0.9%	0.0	0.52	0.6	0.43
	Severe	6.7%	5.9%	0.8	0.45	1.8%	1.9%	-0.1	0.52	0.9	0.42
	Moderate	33.0%	30.8%	2.2	0.45	14.5%	14.8%	-0.3	0.52	2.5	0.40
	Mild	33.1%	33.9%	-0.8	0.45	30.6%	30.9%	-0.3	0.52	-0.5	0.65
	None	23.2%	26.0%	-2.8	0.45	52.2%	51.5%	0.7	0.52	-3.5	0.37
How much difficulty do you have getting on/off toilet?	Extreme	3.0%	1.8%	1.2	0.04	0.6%	0.6%	0.0	0.60	1.2	0.04
	Severe	4.9%	3.2%	1.7	0.04	1.3%	1.2%	0.0	0.60	1.7	0.04
	Moderate	26.7%	19.6%	7.0	0.03	10.3%	10.1%	0.2	0.60	6.8	0.03
	Mild	31.8%	30.3%	1.5	0.14	23.2%	22.9%	0.3	0.60	1.2	0.29
	None	33.6%	45.0%	-11.4	0.03	64.6%	65.2%	-0.5	0.60	-10.9	0.04

Source: CJR evaluation team analysis of patient survey data for episodes with inpatient discharge or outpatient procedure July, August, September, or October 2022.

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.

CJR = Comprehensive Care for Joint Replacement; OTC = over the counter; pp = percentage point.

Exhibit E-18: Risk-Adjusted Survey-Based Results for Differences in Functional Status Between Patients Who Are Black or African American and Patients Who Are Non-Hispanic White Discharged From Mandatory Hospitals

Survey question:	Survey response	Black or African American				Non-Hispanic White				Differential effect	p-value
		CJR risk-adjusted rates (%)	Control risk-adjusted rates (%)	Difference (pp)	p-value	CJR risk-adjusted rates (%)	Control risk-adjusted rates (%)	Difference (pp)	p-value		
How much does pain in the joint interfere with your normal activities?	All of the time	3.4%	3.5%	-0.1	0.94	1.6%	1.6%	0.1	0.48	-0.1	0.89
	Most of the time	7.9%	8.0%	-0.1	0.94	4.2%	4.1%	0.1	0.48	-0.3	0.88
	Some of the time	24.0%	24.3%	-0.2	0.94	16.4%	16.0%	0.4	0.48	-0.6	0.84
	A little of the time	35.0%	35.0%	0.0	0.94	34.6%	34.3%	0.2	0.48	-0.2	0.72
	None of the time	29.7%	29.3%	0.4	0.94	43.2%	44.1%	-0.8	0.48	1.24	0.82
Are you taking any of the following types of medications for the joint you had replaced?	Prescription and OTC	16.0%	17.6%	-1.6	0.65	8.2%	8.8%	-0.6	0.17	-1.0	0.77
	Prescription only	9.8%	10.4%	-0.6	0.65	6.1%	6.5%	-0.3	0.16	-0.2	0.86
	OTC only	29.2%	29.7%	-0.4	0.66	24.3%	25.1%	-0.8	0.16	0.3	0.79
	None	45.0%	42.4%	2.6	0.65	61.4%	59.7%	1.7	0.16	0.9	0.87
What best describes your use of a mobility aid such as a wheelchair, scooter, walker, or cane?	Always Used	23.0%	29.5%	-6.6	0.08	13.5%	13.4%	0.1	0.83	-6.7	0.08
	Sometimes Used	28.6%	30.1%	-1.5	0.13	22.5%	22.4%	0.1	0.83	-1.6	0.15
	Never Used	48.5%	40.4%	8.1	0.09	64.0%	64.3%	-0.2	0.83	8.3	0.09
What best describes your ability to walk by yourself without resting?	Unable to walk by myself without resting	10.5%	13.3%	-2.8	0.17	5.5%	5.7%	-0.2	0.27	-2.5	0.21
	Walk from one room to another by myself	24.6%	27.2%	-2.6	0.17	15.6%	16.1%	-0.4	0.27	-2.1	0.27
	Walk one block by myself without resting	21.2%	21.2%	0.0	0.96	19.0%	19.3%	-0.3	0.27	0.30	0.34
	Walk several blocks by myself without resting	20.3%	19.0%	1.3	0.17	23.5%	23.5%	0.0	0.30	1.3	0.18
	Walk more than several blocks by myself without resting	23.3%	19.3%	4.0	0.17	36.3%	35.4%	0.9	0.27	3.1	0.31

Survey question:	Survey response	Black or African American				Non-Hispanic White				Differential effect	p-value
		CJR risk-adjusted rates (%)	Control risk-adjusted rates (%)	Difference (pp)	p-value	CJR risk-adjusted rates (%)	Control risk-adjusted rates (%)	Difference (pp)	p-value		
How much difficulty do you have walking up or down 12 stairs?	I am not able to	13.6%	15.5%	-1.9	0.51	8.7%	8.7%	-0.1	0.88	-1.8	0.52
	A lot of difficulty	11.6%	12.6%	-1.1	0.50	8.4%	8.4%	0.0	0.88	-1.0	0.52
	Some difficulty	48.0%	48.1%	-0.1	0.78	46.2%	46.3%	-0.1	0.88	0.0	0.97
	No difficulty	26.8%	23.8%	3.0	0.50	36.7%	36.5%	0.2	0.88	2.9	0.54
How much difficulty do you have rising from sitting?	Extreme	1.5%	2.2%	-0.7	0.17	1.0%	1.0%	0.0	0.54	-0.7	0.19
	Severe	3.3%	4.7%	-1.4	0.16	2.3%	2.3%	-0.1	0.55	-1.3	0.19
	Moderate	25.1%	30.7%	-5.6	0.15	19.2%	19.6%	-0.4	0.54	-5.2	0.19
	Mild	37.0%	36.4%	0.6	0.42	36.4%	36.6%	-0.2	0.54	0.8	0.33
	None	33.0%	26.0%	7.1	0.15	41.1%	40.4%	0.7	0.54	6.4	0.21
How much difficulty do you have standing?	Extreme	1.5%	2.0%	-0.5	0.23	0.8%	0.8%	0.0	0.72	-0.5	0.25
	Severe	3.1%	4.2%	-1.0	0.22	1.8%	1.8%	0.0	0.72	-1.0	0.24
	Moderate	21.5%	25.7%	-4.2	0.21	14.3%	14.5%	-0.2	0.72	-4.0	0.23
	Mild	33.9%	34.5%	-0.7	0.29	30.3%	30.4%	-0.2	0.72	-0.5	0.52
	None	40.1%	33.6%	6.4	0.20	52.8%	52.3%	0.4	0.72	6.0	0.25
How much difficulty do you have getting on/off toilet?	Extreme	0.9%	1.1%	-0.2	0.44	0.6%	0.6%	0.0	0.23	-0.2	0.35
	Severe	1.9%	2.3%	-0.4	0.44	1.3%	1.2%	0.1	0.23	-0.4	0.35
	Moderate	14.7%	16.7%	-2.0	0.44	10.5%	9.9%	0.6	0.22	-2.6	0.33
	Mild	27.4%	28.8%	-1.4	0.45	23.2%	22.5%	0.7	0.23	-2.1	0.27
	None	55.1%	51.2%	3.9	0.44	64.4%	65.8%	-1.4	0.23	5.3	0.31

Source: CJR evaluation team analysis of patient survey data for episodes with inpatient discharge or outpatient procedure July, August, September, or October 2022.

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.

CJR = Comprehensive Care for Joint Replacement; OTC = over the counter; pp = percentage point.

Exhibit E-19: Risk-Adjusted Survey-Based Results for Differences in Functional Status Between Patients Who Are Hispanic and Patients Who Are Non-Hispanic White Discharged From Mandatory Hospitals

Survey question:	Survey response	Hispanic				Non-Hispanic White				Differential effect	p-value
		CJR risk-adjusted rates (%)	Control risk-adjusted rates (%)	Difference (pp)	p-value	CJR risk-adjusted rates (%)	Control risk-adjusted rates (%)	Difference (pp)	p-value		
How much does pain in the joint interfere with your normal activities?	All of the time	3.8%	3.3%	0.5	0.62	1.6%	1.6%	0.1	0.48	0.4	0.66
	Most of the time	8.8%	7.9%	0.9	0.61	4.2%	4.1%	0.1	0.48	0.8	0.67
	Some of the time	26.4%	24.8%	1.6	0.61	16.4%	16.0%	0.4	0.48	1.2	0.71
	A little of the time	35.3%	35.7%	-0.5	0.63	34.6%	34.3%	0.2	0.48	-0.7	0.49
	None of the time	25.7%	28.3%	-2.6	0.61	43.2%	44.1%	-0.8	0.48	-1.7	0.74
Are you taking any of the following types of medications for the joint you had replaced?	Prescription and OTC	19.7%	13.4%	6.3	0.07	8.2%	8.8%	-0.6	0.17	6.9	0.05
	Prescription only	11.1%	8.7%	2.4	0.06	6.1%	6.5%	-0.3	0.16	2.7	0.04
	OTC only	30.4%	28.4%	2.0	0.10	24.3%	25.1%	-0.8	0.16	2.8	0.04
	None	38.8%	49.5%	-10.7	0.06	61.4%	59.7%	1.7	0.16	-12.4	0.04
What best describes your use of a mobility aid such as a wheelchair, scooter, walker, or cane?	Always Used	32.0%	32.9%	-0.9	0.81	13.5%	13.4%	0.1	0.83	-1.0	0.79
	Sometimes Used	27.7%	27.8%	-0.2	0.81	22.5%	22.4%	0.1	0.83	-0.3	0.75
	Never Used	40.3%	39.3%	1.1	0.81	64.0%	64.3%	-0.2	0.83	1.3	0.78
What best describes your ability to walk by yourself without resting?	Unable to walk by myself without resting	11.2%	11.8%	-0.7	0.72	5.5%	5.7%	-0.2	0.27	-0.4	0.81
	Walk from one room to another by myself	21.2%	22.0%	-0.8	0.72	15.6%	16.1%	-0.4	0.27	-0.3	0.88
	Walk one block by myself without resting	20.1%	20.3%	-0.2	0.73	19.0%	19.3%	-0.3	0.27	0.1	0.89
	Walk several blocks by myself without resting	20.9%	20.6%	0.3	0.72	23.5%	23.5%	0.0	0.30	0.3	0.75
	Walk more than several blocks by myself without resting	26.6%	25.2%	1.4	0.72	36.3%	35.4%	0.9	0.27	0.5	0.91

Survey question:	Survey response	Hispanic				Non-Hispanic White				Differential effect	p-value
		CJR risk-adjusted rates (%)	Control risk-adjusted rates (%)	Difference (pp)	p-value	CJR risk-adjusted rates (%)	Control risk-adjusted rates (%)	Difference (pp)	p-value		
How much difficulty do you have walking up or down 12 stairs?	I was not able to	17.8%	18.7%	-0.9	0.75	8.7%	8.7%	-0.1	0.88	-0.8	0.77
	A lot of difficulty	12.1%	12.5%	-0.4	0.75	8.4%	8.4%	0.0	0.88	-0.4	0.78
	Some difficulty	46.6%	46.6%	0.0	0.93	46.2%	46.3%	-0.1	0.88	0.1	0.86
	No difficulty	23.5%	22.2%	1.3	0.75	36.7%	36.5%	0.2	0.88	1.1	0.79
How much difficulty do you have rising from sitting?	Extreme	3.0%	2.9%	0.0	0.98	1.0%	1.0%	0.0	0.54	0.1	0.94
	Severe	5.2%	5.2%	0.0	0.98	2.3%	2.3%	-0.1	0.55	0.1	0.92
	Moderate	29.6%	29.5%	0.1	0.98	19.2%	19.6%	-0.4	0.54	0.5	0.89
	Mild	35.8%	35.9%	0.0	0.98	36.4%	36.6%	-0.2	0.54	0.2	0.82
	None	26.4%	26.6%	-0.1	0.98	41.1%	40.4%	0.7	0.54	-0.8	0.86
How much difficulty do you have standing?	Extreme	3.6%	3.2%	0.4	0.63	0.8%	0.8%	0.0	0.72	0.4	0.62
	Severe	5.7%	5.2%	0.5	0.63	1.8%	1.8%	0.0	0.72	0.5	0.61
	Moderate	27.9%	26.3%	1.6	0.63	14.3%	14.5%	-0.2	0.72	1.8	0.60
	Mild	32.8%	32.8%	-0.1	0.75	30.3%	30.4%	-0.2	0.72	0.1	0.85
	None	30.0%	32.4%	-2.4	0.63	52.8%	52.3%	0.4	0.72	-2.8	0.59
How much difficulty do you have getting on/off toilet?	Extreme	2.7%	2.2%	0.5	0.40	0.6%	0.6%	0.0	0.23	0.5	0.44
	Severe	3.9%	3.3%	0.6	0.40	1.3%	1.2%	0.1	0.23	0.6	0.46
	Moderate	21.1%	18.4%	2.7	0.39	10.5%	9.9%	0.6	0.22	2.2	0.50
	Mild	29.1%	28.0%	1.2	0.39	23.2%	22.5%	0.7	0.23	0.4	0.77
	None	43.2%	48.2%	-5.0	0.39	64.4%	65.8%	-1.4	0.23	-3.6	0.54

Source: CJR evaluation team analysis of patient survey data for episodes with inpatient discharge or outpatient procedure July, August, September, or October 2022.

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.

CJR = Comprehensive Care for Joint Replacement; OTC = over the counter; pp = percentage point.

2. Satisfaction With Care Management and Overall Recovery

Exhibit E-20: Risk-Adjusted Survey-Based Results for Differences in Satisfaction Between Dually Eligible Patients and Patients Who Were Not Dually Eligible Discharged From Mandatory Hospitals

Survey question:	Survey response	Dually eligible				Non-dually eligible				Differential effect	p-value
		CJR risk-adjusted rates (%)	Control risk-adjusted rates (%)	Difference (pp)	p-value	CJR risk-adjusted rates (%)	Control risk-adjusted rates (%)	Difference (pp)	p-value		
How satisfied or dissatisfied were you with your overall recovery since you left the hospital?	Very dissatisfied	9.4%	10.2%	-0.8	0.75	7.1%	7.3%	-0.2	0.59	-0.7	0.80
	Dissatisfied	7.3%	7.8%	-0.5	0.74	5.9%	6.0%	-0.1	0.59	-0.4	0.81
	Neither	3.4%	3.6%	-0.2	0.74	2.8%	2.9%	-0.1	0.59	-0.1	0.81
	Somewhat satisfied	22.0%	22.7%	-0.7	0.74	19.7%	19.9%	-0.2	0.59	-0.5	0.83
	Very satisfied	57.9%	55.7%	2.3	0.74	64.5%	63.9%	0.6	0.59	1.6	0.81
How satisfied or dissatisfied were you with the extent to which healthcare providers listened to your thoughts and preferences about treatment?	Very dissatisfied	16.2%	17.6%	-1.4	0.68	11.7%	11.9%	-0.2	0.67	-1.2	0.73
	Dissatisfied	5.7%	6.1%	-0.4	0.68	4.5%	4.6%	-0.1	0.67	-0.3	0.75
	Neither	5.1%	5.4%	-0.3	0.68	4.1%	4.2%	-0.1	0.67	-0.2	0.75
	Somewhat satisfied	13.6%	14.0%	-0.4	0.68	11.8%	11.9%	-0.1	0.67	-0.3	0.78
	Very satisfied	59.3%	56.9%	2.4	0.68	67.8%	67.4%	0.5	0.67	2.0	0.74
How satisfied or dissatisfied were you with the place you were sent after you left the hospital?	Very dissatisfied	16.8%	23.7%	-6.9	0.09	9.7%	10.0%	-0.3	0.55	-6.6	0.11
	Dissatisfied	4.5%	5.6%	-1.1	0.08	3.0%	3.0%	-0.1	0.55	-1.1	0.10
	Neither	3.8%	4.6%	-0.8	0.08	2.7%	2.7%	-0.1	0.56	-0.7	0.11
	Somewhat satisfied	8.7%	9.9%	-1.2	0.08	6.5%	6.6%	-0.1	0.55	-1.1	0.13
	Very satisfied	66.3%	56.2%	10.1	0.08	78.1%	77.6%	0.6	0.55	9.5	0.11

Survey question:	Survey response	Dually eligible				Non-dually eligible				Differential effect	p-value
		CJR risk-adjusted rates (%)	Control risk-adjusted rates (%)	Difference (pp)	p-value	CJR risk-adjusted rates (%)	Control risk-adjusted rates (%)	Difference (pp)	p-value		
How satisfied or dissatisfied were you with the coordination of your care among doctors, nurses, and therapists in the hospital and after discharge?	Very dissatisfied	14.2%	15.4%	-1.2	0.71	8.4%	8.5%	-0.1	0.80	-1.1	0.74
	Dissatisfied	5.8%	6.1%	-0.3	0.71	3.8%	3.9%	0.0	0.80	-0.3	0.75
	Neither	4.0%	4.2%	-0.2	0.71	2.8%	2.8%	0.0	0.80	-0.2	0.75
	Somewhat satisfied	13.3%	13.7%	-0.4	0.71	10.2%	10.3%	-0.1	0.80	-0.4	0.77
	Very satisfied	62.7%	60.6%	2.1	0.71	74.8%	74.5%	0.3	0.80	1.9	0.75
How satisfied or dissatisfied were you with the instructions you received from doctors, nurses, and therapists about your treatment?	Very dissatisfied	9.9%	11.0%	-1.2	0.61	7.1%	7.1%	0.0	0.99	-1.2	0.62
	Dissatisfied	4.7%	5.2%	-0.4	0.61	3.6%	3.6%	0.0	0.99	-0.4	0.62
	Neither	4.2%	4.5%	-0.3	0.61	3.3%	3.3%	0.0	0.99	-0.3	0.62
	Somewhat satisfied	16.1%	16.9%	-0.9	0.61	13.5%	13.5%	0.0	0.99	-0.9	0.62
	Very satisfied	65.1%	62.4%	2.8	0.61	72.6%	72.6%	0.0	0.99	2.8	0.62

Source: CJR evaluation team analysis of patient survey data for episodes with inpatient discharge or outpatient procedure July, August, September, or October 2022.

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.

CJR = Comprehensive Care for Joint Replacement; pp = percentage point.

Exhibit E-21: Risk-Adjusted Survey-Based Results for Differences in Satisfaction Between Patients Who Are Black or African American and Patients Who Are Non-Hispanic White Discharged From Mandatory Hospitals

Survey question:	Survey response	Black or African American				Non-Hispanic White				Differential effect	p-value
		CJR risk-adjusted rates (%)	Control risk-adjusted rates (%)	Difference (pp)	p-value	CJR risk-adjusted rates (%)	Control risk-adjusted rates (%)	Difference (pp)	p-value		
How satisfied or dissatisfied were you with your overall recovery since you left the hospital?	Very dissatisfied	7.5%	9.5%	-2.0	0.32	7.0%	7.1%	0.0	0.90	-1.9	0.34
	Dissatisfied	6.1%	7.4%	-1.3	0.32	5.8%	5.8%	0.0	0.90	-1.3	0.34
	Neither	2.9%	3.4%	-0.5	0.32	2.8%	2.8%	0.0	0.90	-0.5	0.34
	Somewhat satisfied	20.1%	22.3%	-2.2	0.31	19.5%	19.6%	-0.1	0.90	-2.1	0.34
	Very satisfied	63.3%	57.4%	6.0	0.31	64.9%	64.7%	0.2	0.90	5.8	0.34
How satisfied or dissatisfied were you with the extent to which healthcare providers listened to your thoughts and preferences about treatment?	Very dissatisfied	12.0%	16.8%	-4.7	0.10	11.6%	11.8%	-0.1	0.82	-4.6	0.12
	Dissatisfied	4.6%	5.9%	-1.3	0.10	4.5%	4.5%	0.0	0.82	-1.3	0.12
	Neither	4.2%	5.2%	-1.0	0.10	4.1%	4.2%	0.0	0.82	-1.0	0.12
	Somewhat satisfied	12.0%	13.9%	-1.9	0.10	11.8%	11.8%	-0.1	0.81	-1.8	0.13
	Very satisfied	67.1%	58.2%	8.9	0.10	68.0%	67.7%	0.3	0.82	8.6	0.12
How satisfied or dissatisfied were you with the place you were sent after you left the hospital?	Very dissatisfied	12.5%	17.4%	-5.0	0.13	9.6%	9.7%	-0.2	0.78	-4.8	0.15
	Dissatisfied	3.6%	4.6%	-1.0	0.13	2.9%	3.0%	0.0	0.78	-1.0	0.16
	Neither	3.2%	4.0%	-0.8	0.13	2.6%	2.7%	0.0	0.78	-0.7	0.16
	Somewhat satisfied	7.6%	9.0%	-1.4	0.14	6.4%	6.5%	-0.1	0.78	-1.4	0.17
	Very satisfied	73.2%	65.0%	8.2	0.13	78.4%	78.1%	0.3	0.78	7.9	0.15
How satisfied or dissatisfied were you with the coordination of your care among doctors, nurses, and therapists in the hospital and after discharge?	Very dissatisfied	10.4%	12.6%	-2.1	0.40	8.2%	8.4%	-0.1	0.78	-2.0	0.44
	Dissatisfied	4.6%	5.3%	-0.7	0.40	3.8%	3.8%	-0.1	0.78	-0.7	0.44
	Neither	3.3%	3.7%	-0.4	0.40	2.7%	2.8%	<0.1	0.78	-0.4	0.45
	Somewhat satisfied	11.5%	12.7%	-1.2	0.40	10.1%	10.2%	-0.1	0.78	-1.1	0.45
	Very satisfied	70.2%	65.7%	4.5	0.40	75.2%	74.9%	0.3	0.78	4.2	0.44

Survey question:	Survey response	Black or African American				Non-Hispanic White				Differential effect	p-value
		CJR risk-adjusted rates (%)	Control risk-adjusted rates (%)	Difference (pp)	p-value	CJR risk-adjusted rates (%)	Control risk-adjusted rates (%)	Difference (pp)	p-value		
How satisfied or dissatisfied were you with the instructions you received from doctors, nurses, and therapists about your treatment?	Very dissatisfied	8.4%	9.2%	-0.7	0.75	6.9%	6.8%	0.1	0.70	-0.9	0.71
	Dissatisfied	4.2%	4.5%	-0.3	0.75	3.5%	3.5%	0.1	0.70	-0.4	0.70
	Neither	3.8%	4.0%	-0.2	0.75	3.2%	3.2%	0.1	0.70	-0.3	0.70
	Somewhat satisfied	15.1%	15.8%	-0.7	0.75	13.4%	13.2%	0.2	0.70	-0.9	0.69
	Very satisfied	68.5%	66.5%	2.0	0.75	73.0%	73.5%	-0.4	0.70	2.4	0.70

Source: CJR evaluation team analysis of patient survey data for episodes with inpatient discharge or outpatient procedure July, August, September, or October 2022.

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.

CJR = Comprehensive Care for Joint Replacement; pp = percentage point.

Exhibit E-22: Risk-Adjusted Survey-Based Results for Differences in Satisfaction Between Patients Who Are Hispanic and Patients Who Are Non-Hispanic White Discharged From Mandatory Hospitals

Survey question:	Survey response	Hispanic				Non-Hispanic White				Differential effect	p-value
		CJR risk-adjusted rates (%)	Control risk-adjusted rates (%)	Difference (pp)	p-value	CJR risk-adjusted rates (%)	Control risk-adjusted rates (%)	Difference (pp)	p-value		
How satisfied or dissatisfied were you with your overall recovery since you left the hospital?	Very dissatisfied	8.7%	9.3%	-0.6	0.78	7.0%	7.1%	0.0	0.90	-0.6	0.80
	Dissatisfied	6.9%	7.3%	-0.4	0.78	5.8%	5.8%	0.0	0.90	-0.4	0.80
	Neither	3.2%	3.4%	-0.2	0.78	2.8%	2.8%	0.0	0.90	-0.1	0.80
	Somewhat satisfied	21.3%	21.9%	-0.6	0.78	19.5%	19.6%	-0.1	0.90	-0.6	0.81
	Very satisfied	60.0%	58.1%	1.9	0.78	64.9%	64.7%	0.2	0.90	1.7	0.80
How satisfied or dissatisfied were you with the extent to which healthcare providers listened to your thoughts and preferences about treatment?	Very dissatisfied	12.3%	14.8%	-2.5	0.50	11.6%	11.8%	-0.1	0.82	-2.3	0.53
	Dissatisfied	4.6%	5.3%	-0.7	0.50	4.5%	4.5%	0.0	0.82	-0.6	0.53
	Neither	4.2%	4.8%	-0.5	0.50	4.1%	4.2%	0.0	0.82	-0.5	0.53
	Somewhat satisfied	11.9%	12.9%	-1.0	0.49	11.8%	11.8%	-0.1	0.81	-1.0	0.53
	Very satisfied	67.0%	62.2%	4.7	0.50	68.0%	67.7%	0.3	0.82	4.5	0.53
How satisfied or dissatisfied were you with the place you were sent after you left the hospital?	Very dissatisfied	14.4%	19.0%	-4.6	0.24	9.6%	9.7%	-0.2	0.78	-4.5	0.26
	Dissatisfied	3.8%	4.6%	-0.8	0.23	2.9%	3.0%	<0.1	0.78	-0.8	0.26
	Neither	3.2%	3.8%	-0.6	0.22	2.6%	2.7%	<0.1	0.78	-0.6	0.26
	Somewhat satisfied	7.5%	8.6%	-1.1	0.22	6.4%	6.5%	-0.1	0.78	-1.0	0.27
	Very satisfied	71.1%	64.0%	7.2	0.23	78.4%	78.1%	0.3	0.78	6.9	0.26
How satisfied or dissatisfied were you with the coordination of your care among doctors, nurses, and therapists in the hospital and after discharge?	Very dissatisfied	10.3%	12.4%	-2.1	0.51	8.2%	8.4%	-0.1	0.78	-2.0	0.54
	Dissatisfied	4.4%	5.1%	-0.7	0.51	3.8%	3.8%	-0.1	0.78	-0.6	0.54
	Neither	3.1%	3.6%	-0.4	0.51	2.7%	2.8%	<0.1	0.78	-0.4	0.55
	Somewhat satisfied	11.0%	12.1%	-1.1	0.50	10.1%	10.2%	-0.1	0.78	-1.0	0.55
	Very satisfied	71.1%	66.8%	4.4	0.51	75.2%	74.9%	0.3	0.78	4.1	0.54

Survey question:	Survey response	Hispanic				Non-Hispanic White				Differential effect	p-value
		CJR risk-adjusted rates (%)	Control risk-adjusted rates (%)	Difference (pp)	p-value	CJR risk-adjusted rates (%)	Control risk-adjusted rates (%)	Difference (pp)	p-value		
How satisfied or dissatisfied were you with the instructions you received from doctors, nurses, and therapists about your treatment?	Very dissatisfied	7.6%	10.2%	-2.5	0.30	6.9%	6.8%	0.1	0.70	-2.7	0.27
	Dissatisfied	3.8%	4.8%	-1.0	0.29	3.5%	3.5%	0.1	0.70	-1.1	0.27
	Neither	3.4%	4.2%	-0.8	0.28	3.2%	3.2%	0.1	0.70	-0.9	0.26
	Somewhat satisfied	13.7%	15.9%	-2.3	0.28	13.4%	13.2%	0.2	0.70	-2.5	0.25
	Very satisfied	71.6%	64.9%	6.6	0.29	73.0%	73.5%	-0.4	0.70	7.1	0.26

Source: CJR evaluation team analysis of patient survey data for episodes with inpatient discharge or outpatient procedure July, August, September, or October 2022.

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.

CJR = Comprehensive Care for Joint Replacement; pp = percentage point.

3. Need for Caregiver Help

Exhibit E-23: Risk-Adjusted Survey-Based Results for Differences in Caregiver Help Between Patients Who Were Dually Eligible and Patients Who Were Not Dually Eligible Discharged From Mandatory Hospitals

Survey question:	Survey response	Dually eligible				Non-dually eligible				Differential effect	p-value
		CJR risk-adjusted rates (%)	Control risk-adjusted rates (%)	Difference (pp)	p-value	CJR risk-adjusted rates (%)	Control risk-adjusted rates (%)	Difference (pp)	p-value		
How much caregiver help did you need putting on or taking off clothes?	Complete help needed	18.6%	14.9%	3.7	0.29	10.0%	9.7%	0.3	0.47	3.4	0.34
	Some help needed	58.1%	56.9%	1.2	0.39	56.8%	56.4%	0.4	0.47	0.8	0.61
	No help needed	23.3%	28.2%	-4.8	0.31	33.1%	33.8%	-0.7	0.47	-4.1	0.40
How much caregiver help did you need using the toilet?	Complete help needed	15.1%	14.1%	1.0	0.74	5.1%	5.3%	-0.1	0.67	1.1	0.71
	Some help needed	39.4%	38.4%	0.9	0.74	25.1%	25.5%	-0.3	0.67	1.3	0.66
	No help needed	45.5%	47.5%	-1.9	0.74	69.7%	69.3%	0.5	0.67	-2.4	0.68
How much caregiver help did you need with bathing?	Complete help needed	22.4%	17.2%	5.2	0.19	10.2%	10.2%	0.0	0.95	5.2	0.19
	Some help needed	48.5%	46.6%	1.9	0.28	43.0%	43.1%	0.0	0.95	2.0	0.30
	No help needed	29.1%	36.2%	-7.1	0.20	46.8%	46.7%	0.1	0.95	-7.2	0.20

Source: CJR evaluation team analysis of patient survey data for episodes with inpatient discharge or outpatient procedure July, August, September, or October 2022.

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.

CJR = Comprehensive Care for Joint Replacement; pp = percentage point.

Exhibit E-24: Risk-Adjusted Survey-Based Results for Differences in Caregiver Help Between Patients Who Are Black or African American and Patients Who Are Non-Hispanic White Discharged From Mandatory Hospitals

Survey question:	Survey response	Black or African American				Non-Hispanic White				Differential effect	p-value
		CJR risk-adjusted rates (%)	Control risk-adjusted rates (%)	Difference (pp)	p-value	CJR risk-adjusted rates (%)	Control risk-adjusted rates (%)	Difference (pp)	p-value		
How much caregiver help did you need putting on or taking off clothes?	Complete help needed	12.7%	12.1%	0.6	0.81	9.8%	9.3%	0.4	0.33	0.2	0.94
	Some help needed	59.1%	58.5%	0.5	0.81	56.5%	55.9%	0.6	0.33	-0.1	0.97
	No help needed	28.2%	29.4%	-1.2	0.81	33.7%	34.8%	-1.1	0.33	-0.1	0.98
How much caregiver help did you need using the toilet?	Complete help needed	8.3%	8.6%	-0.3	0.88	4.9%	5.0%	-0.1	0.80	-0.2	0.91
	Some help needed	32.8%	33.4%	-0.6	0.88	24.4%	24.6%	-0.2	0.80	-0.4	0.93
	No help needed	58.9%	58.0%	0.9	0.88	70.7%	70.4%	0.3	0.80	0.6	0.92
How much caregiver help did you need with bathing?	Complete help needed	15.5%	15.1%	0.3	0.92	9.9%	9.7%	0.2	0.61	0.1	0.97
	Some help needed	48.8%	48.5%	0.3	0.92	42.7%	42.3%	0.4	0.61	-0.1	0.96
	No help needed	35.7%	36.3%	-0.6	0.92	47.4%	48.0%	-0.6	0.61	0.0	1.00

Source: CJR evaluation team analysis of patient survey data for episodes with inpatient discharge or outpatient procedure July, August, September, or October 2022.

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.

CJR = Comprehensive Care for Joint Replacement; pp = percentage point.

Exhibit E-25: Risk-Adjusted Survey-Based Results for Differences in Caregiver Help Between Patients Who Are Hispanic and Patients Who Are Non-Hispanic White Discharged From Mandatory Hospitals

Survey question:	Survey response	Hispanic				Non-Hispanic White				Differential effect	p-value
		CJR risk-adjusted rates (%)	Control risk-adjusted rates (%)	Difference (pp)	p-value	CJR risk-adjusted rates (%)	Control risk-adjusted rates (%)	Difference (pp)	p-value		
How much caregiver help did you need putting on or taking off clothes?	Complete help needed	23.8%	16.4%	7.4	0.08	9.8%	9.3%	0.4	0.33	7.0	0.10
	Some help needed	59.9%	59.7%	0.2	0.84	56.5%	55.9%	0.6	0.33	-0.4	0.76
	No help needed	16.3%	23.9%	-7.6	0.08	33.7%	34.8%	-1.1	0.33	-6.6	0.15
How much caregiver help did you need using the toilet?	Complete help needed	17.0%	15.3%	1.7	0.63	4.9%	5.0%	-0.1	0.80	1.8	0.62
	Some help needed	40.3%	38.9%	1.5	0.64	24.4%	24.6%	-0.2	0.80	1.7	0.60
	No help needed	42.7%	45.8%	-3.2	0.63	70.7%	70.4%	0.3	0.80	-3.5	0.61
How much caregiver help did you need with bathing?	Complete help needed	23.5%	23.8%	-0.3	0.95	9.9%	9.7%	0.2	0.61	-0.5	0.92
	Some help needed	50.2%	50.2%	0.0	0.95	42.7%	42.3%	0.4	0.61	-0.4	0.67
	No help needed	26.3%	25.9%	0.3	0.95	47.4%	48.0%	-0.6	0.61	0.9	0.87

Source: CJR evaluation team analysis of patient survey data for episodes with inpatient discharge or outpatient procedure July, August, September, or October 2022.

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.

CJR = Comprehensive Care for Joint Replacement; pp = percentage point.

Appendix F: Claims-Based Health Equity Analyses

A. Extended Claims-Based Outcomes

Exhibit F-1A: The CJR Model Had a Statistically Significant Impact on LEJR Volume for Beneficiaries Who Are Hispanic

Population	CJR			Control			Impact (DiD)	p-value	90% LCI	90% UCI
	PY 6 bene-years (N)	Baseline mean	PY 6 mean	PY 6 bene-years (N)	Baseline mean	PY 6 mean				
Beneficiaries who are Black or African American	318,394	686	602	348,949	684	611	-11	0.77	-72	50
Beneficiaries who are Non-Hispanic White	3,057,431	1,194	1,159	3,182,114	1,239	1,159	45	0.15	-6	95
Beneficiaries who were dually eligible	749,748	532	411	509,465	590	429	41	0.11	-1	82
Beneficiaries who were not dually eligible	3,330,992	1,218	1,198	3,427,988	1,252	1,189	42	0.15	-6	90
Beneficiaries who are Black or African American and dually eligible	112,561	513	369	113,879	539	397	-1	0.98	-76	74
Beneficiaries who are Non-Hispanic White and not dually eligible	3,125,159	1,250	1,229	3,192,918	1,297	1,230	46	0.11	-2	94
Beneficiaries who are Hispanic	331,444	645	603	188,029	685	571	71	<0.01	28	115
Beneficiaries who are Non-Hispanic White	3,057,431	1,194	1,160	3,182,114	1,233	1,156	43	0.16	-7	94

Source: CJR evaluation team analysis of Medicare claims and enrollment data for Medicare FFS beneficiaries enrolled in Medicare between 2012 and 2015 (baseline) or during 2022 (intervention).

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

CJR = Comprehensive Care for Joint Replacement; DiD = difference-in-differences; DDD = difference-in-difference-in-differences; LCI = lower confidence interval; LEJR = lower extremity joint replacement; PY = performance year; UCI = upper confidence interval.

[a] The estimate for this reference population differs depending on the underserved population because each underserved-reference population pair is handled in a separate regression.

Exhibit F-1B: The CJR Model Had a Statistically Significant Impact on LEJR Volume for Beneficiaries Who Are Hispanic

Population	Impact	Difference in impact	p-value	90% LCI	90% UCI
Beneficiaries who are Black or African American	-11	-56	0.13	-116	5
Beneficiaries who are Non-Hispanic White [a]	45				
Beneficiaries who were dually eligible	41	-2	0.94	-41	37
Beneficiaries who were not dually eligible	42				
Beneficiaries who are Black or African American and dually eligible	-1	-47	0.36	-132	38
Beneficiaries who are Non-Hispanic White and not dually eligible	46				
Beneficiaries who are Hispanic	71	28	0.40	-26	82
Beneficiaries who are Non-Hispanic White [a]	43				

Source: CJR evaluation team analysis of Medicare claims and enrollment data for Medicare FFS beneficiaries enrolled in Medicare between 2012 and 2015 (baseline) or during 2022 (intervention).

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

CJR = Comprehensive Care for Joint Replacement; DDD = difference-in-difference-in-differences; FFS = fee-for-service; LCI = lower confidence interval; LEJR = lower extremity joint replacement; UCI = upper confidence interval.

[a] The estimate for this reference population differs depending on the underserved population because each underserved-reference population pair is handled in a separate regression.

Exhibit F-2A: The CJR Model Likely Reduced Average Episode Payments for All Patients, but Reduced Average Episode Payments by More for Patients From Underserved Populations

Population	CJR			Control			Impact (DiD)	p-value	90% LCI	90% UCI
	PY 6 episodes (N)	Baseline mean	PY 6 mean	PY 6 episodes (N)	Baseline mean	PY 6 mean				
Patients who are Black or African American	1,943	$28,295	$23,677	2,662	$26,212	$23,725	-$2,132*	0.07	-$4,056	-$207
Patients who are Non-Hispanic White [a]	40,046	$25,329	$21,633	44,909	$24,430	$21,886	-$1,152	0.10	-$2,296	-$8
Patients who were dually eligible	3,103	$30,388	$26,272	2,577	$28,208	$26,100	-$2,008*	0.02	-$3,431	-$585
Patients who were not dually eligible	43,908	$25,106	$21,333	48,502	$24,199	$21,588	-$1,161	0.10	-$2,327	$5
Patients who are Black or African American and dually eligible	396	$30,156	$26,168	483	$27,424	$25,417	-$1,980*	0.24	-$4,777	$817
Patients who are Non-Hispanic White and not dually eligible	38,597	$25,002	$21,347	43,398	$24,106	$21,547	-$1,097	0.11	-$2,236	$42
Patients who are Hispanic	2,312	$27,830	$23,008	1,328	$25,309	$22,678	-$2,191	<0.01	-$3,469	-$913
Patients who are Non-Hispanic White [a]	40,046	$25,382	$21,525	44,909	$24,463	$21,746	-$1,140	0.10	-$2,289	$9

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 ending between October 1, 2021, and December 31, 2022 (PY 6 intervention).

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

CJR = Comprehensive Care for Joint Replacement; DiD = difference-in-differences; DDD = difference-in-difference-in-differences; LCI = lower confidence interval; PY = performance year; UCI = upper confidence interval.

* The sample for this estimate failed tests for parallel trends in the baseline period.

[a] The estimate for this reference population differs depending on the underserved population because each underserved-reference population pair is handled in a separate regression.

Exhibit F-2B: The CJR Model Likely Reduced Average Episode Payments for All Patients but Reduced Average Episode Payments by More for Patients From Underserved Populations

Population	Impact	Difference in impact	p-value	90% LCI	90% UCI
Patients who are Black or African American	-$2,132*	-$980*	0.15	-$2,105	$146
Patients who are Non-Hispanic White [a]	-$1,152				
Patients who were dually eligible	-$2,008*	-$847	0.08	-$1,635	-$58
Patients who were not dually eligible	-$1,161				
Patients who are Black or African American and dually eligible	-$1,980*	-$883	0.47	-$2,921	$1,154
Patients who are Non-Hispanic White and not dually eligible	-$1,097				
Patients who are Hispanic	-$2,191	-$1,051	0.01	-$1,739	-$363
Patients who are Non-Hispanic White [a]	-$1,140				

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 ending between October 1, 2021, and December 31, 2022 (PY 6 intervention).

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

CJR = Comprehensive Care for Joint Replacement; DDD = difference-in-difference-in-differences; LCI = lower confidence interval; PY = performance year; UCI = upper confidence interval.

* The sample for this estimate failed tests for parallel trends in the baseline period.

[a] The estimate for this reference population differs depending on the underserved population because each underserved-reference population pair is handled in a separate regression.

Exhibit F-3A: Patients Who Are Black or African American Had a Statistically Significant Decrease in All-Cause Mortality and a Statistically Significant Increase in ED Use. The Latter Was Significantly Different From the Impact for Patients Who Are Non-Hispanic White

Outcome	Population	CJR			Control			Impact (DiD)	p-value	90% LCI	90% UCI
		PY 6 episodes (N)	Baseline mean	PY 6 mean	PY 6 episodes (N)	Baseline mean	PY 6 mean				
Mortality rate	Patients who are Black or African American	1,940	0.4%	0.2%	2,673	0.5%	0.8%	-0.4 pp	0.08	-0.8	-0.0
	Patients who are Non-Hispanic White	40,160	0.5%	0.5%	45,030	0.5%	0.5%	0.0 pp	0.77	-0.1	0.1
ED use	Patients who are Black or African American	1,943	17.6%	17.1%	2,661	16.8%	14.1%	2.2 pp	0.08	0.1	4.3
	Patients who are Non-Hispanic White	40,035	11.8%	11.9%	44,902	11.2%	11.2%	0.0 pp	0.96	-0.7	0.7
Readmission rate	Patients who are Black or African American	1,941	9.8%	6.9%	2,659	8.9%	6.8%	-0.8 pp	0.38	-2.3	0.7
	Patients who are Non-Hispanic White	40,016	7.9%	5.5%	44,881	7.6%	5.6%	-0.5 pp	0.08	-1.0	-0.0

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 ending between October 1, 2021, and December 31, 2022 (PY 6 intervention).

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

CJR = Comprehensive Care for Joint Replacement; DiD = difference-in-differences; DDD = difference-in-difference-in-differences; ED = emergency department; LCI = lower confidence interval; PY = performance year; UCI = upper confidence interval.

Exhibit F-3B: Patients Who Are Black or African American Experienced a Statistically Significant Decrease in All-Cause Mortality and a Statistically Significant Increase in ED Use. The Latter Was Significantly Different From the Impact for Patients Who Are Non-Hispanic White

Outcome	Impact on:		Difference in impact	p-value	90% LCI	90% UCI
	Patients who are Black or African American	Patients who are Non-Hispanic White				
Mortality rate	-0.4 pp	0.0 pp	-0.4 pp	0.11	-0.8	0.0
ED use	2.2 pp	0.0 pp	2.2 pp	0.09	0.1	4.3
Readmission rate	-0.8 pp	-0.5 pp	-0.3 pp	0.76	-1.7	1.2

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 ending between October 1, 2021, and December 31, 2022 (PY 6 intervention).

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

CJR = Comprehensive Care for Joint Replacement; DDD = difference-in-difference-in-differences; ED = emergency department; LCI = lower confidence interval; PY = performance year; UCI = upper confidence interval.

Exhibit F-4A: The CJR Model Led to a Larger Decrease in First Discharge to IRF for Patients Who Are Black or African American Than for Patients Who Are Non-Hispanic White and a Significantly Different Impact on First Discharge Without Home Health

First PAC discharge	Population	CJR			Control			Impact (DiD)	p-value	90% LCI	90% UCI
		PY 6 episodes (N)	Baseline mean	PY 6 mean	PY 6 episodes (N)	Baseline mean	PY 6 mean				
IRF	Patients who are Black or African American	1,943	16.1%	4.5%	2,662	12.8%	7.7%	-6.5 pp	0.03	-11.3	-1.6
	Patients who are Non-Hispanic White	40,046	10.3%	2.2%	44,909	9.8%	5.0%	-3.3 pp	0.10	-6.6	0.0
SNF	Patients who are Black or African American	1,943	42.4%	13.8%	2,662	38.7%	12.1%	-2.0 pp	0.65	-9.3	5.3
	Patients who are Non-Hispanic White	40,046	37.8%	7.9%	44,909	35.6%	9.0%	-3.4 pp	0.14	-7.1	0.4
HH	Patients who are Black or African American	1,943	33.1%	63.7%	2,662	31.3%	55.0%	7.0 pp	0.42	-7.2	21.1
	Patients who are Non-Hispanic White	40,046	38.7%	71.9%	44,909	35.4%	56.8%	11.8 pp*	0.07	1.0	22.5
No HH	Patients who are Black or African American	1,943	8.4%	18.0%	2,662	17.2%	25.3%	1.5 pp	0.78	-7.1	10.1
	Patients who are Non-Hispanic White	40,046	13.2%	18.1%	44,909	19.2%	29.1%	-5.1 pp	0.28	-12.8	2.7

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 ending between October 1, 2021, and December 31, 2022 (PY 6 intervention).

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

CJR = Comprehensive Care for Joint Replacement; DiD = difference-in-differences; DDD = difference-in-difference-in-differences; HH = home health; IRF = inpatient rehabilitation facility; LCI = lower confidence interval; PY = performance year; SNF = skilled nursing facility; UCI = upper confidence interval.

* The sample for this estimate failed tests for parallel trends in the baseline period.

Exhibit F-4B: Patients Who Are Black or African American Had a Larger Negative Impact on First Discharge to IRF Than Patients Who Are Non-Hispanic White and a Significantly Different Impact on First Discharge Without Home Health

First PAC discharge	Impact on:		Difference in impact	p-value	90% LCI	90% UCI
	Patients who are Black or African American	Patients who are Non-Hispanic White				
IRF	-6.5 pp	-3.3 pp	-3.1 pp	0.09	-6.2	-0.1
SNF	-2.0 pp	-3.4 pp	1.4 pp	0.66	-3.9	6.7
HH	7.0 pp	11.8 pp*	-4.8 pp	0.24	-11.6	2.0
No HH	1.5 pp	-5.1 pp	6.5 pp	<0.01	2.9	10.1

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 ending between October 1, 2021, and December 31, 2022 (PY 6 intervention).

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

CJR = Comprehensive Care for Joint Replacement; DDD = difference-in-difference-in-differences; HH = home health; IRF = inpatient rehabilitation facility; LCI = lower confidence interval; PY = performance year; SNF = skilled nursing facility; UCI = upper confidence interval.

* The sample for this estimate failed tests for parallel trends in the baseline period.

Exhibit F-5A: Patients Who Are Black or African American Had a Statistically Significant Decrease in SNF Length of Stay, but the Difference Relative to Patients Who Are Non-Hispanic White Was Not Statistically Significant

Outcome	Population	CJR			Control			Impact (DiD)	p-value	90% LCI	90% UCI
		PY 6 episodes (N)	Baseline mean	PY 6 mean	PY 6 episodes (N)	Baseline mean	PY 6 mean				
HH visits	Patients who are Black or African American	1,295	18.7	10.5	1,620	18.2	10.9	-1.0	0.10	-2.0	0.0
	Patients who are Non-Hispanic White	28,260	17.3	9.7	25,480	16.8	9.9	-0.7*	0.24	-1.7	0.3
SNF LOS	Patients who are Black or African American	307	24.0	19.0	293	22.2	20.5	-3.4*	0.06	-6.3	-0.5
	Patients who are Non-Hispanic White	3,799	19.7	15.1	3,852	20.1	16.1	-0.6	0.49	-2.0	0.8
IRF LOS	Patients who are Black or African American	101	10.5	10.9	164	10.5	11.3	-0.4	0.65	-1.8	1.0
	Patients who are Non-Hispanic White	1,096	10.2	11.2	1,682	10.1	10.7	0.4	0.16	-0.1	0.8

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 ending between October 1, 2021, and December 31, 2022 (PY 6 intervention).

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

CJR = Comprehensive Care for Joint Replacement; DiD = difference-in-differences; DDD = difference-in-difference-in-differences; HH = home health; IRF = inpatient rehabilitation facility; LCI = lower confidence interval; LOS = length of stay; PY = performance year; SNF = skilled nursing facility; UCI = upper confidence interval.

* The sample for this estimate failed tests for parallel trends in the baseline period.

Exhibit F-5B: Patients Who Are Black or African American Had a Statistically Significant Decrease in SNF Length of Stay, but the Difference Relative to Patients Who Are Non-Hispanic White Was Not Statistically Significant

Outcome	Impact on:		Difference in Impact	p-value	90% LCI	90% UCI
	Patients who are Black or African American	Patients who are Non-Hispanic White				
HH visits	-1.0	-0.7*	-0.3	0.54	-1.1	0.5
SNF LOS	-3.4*	-0.6	-2.8*	0.11	-5.7	0.0
IRF LOS	-0.4	0.4	-0.8	0.35	-2.1	0.6

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 ending between October 1, 2021, and December 31, 2022 (PY 6 intervention).

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

CJR = Comprehensive Care for Joint Replacement; DDD = difference-in-difference-in-differences; HH = home health; IRF = inpatient rehabilitation facility; LCI = lower confidence interval; LOS = length of stay; PY = performance year; SNF = skilled nursing facility; UCI = upper confidence interval.

* The sample for this estimate failed tests for parallel trends in the baseline period.

Exhibit F-6A: The Model Did Not Have Any Significant Impacts on Quality Measures for Patients Who Were or Were Not Dually Eligible, and Differences Between the Populations Were Not Statistically Significant

Outcome	Population	CJR			Control			Impact (DiD)	p-value	90% LCI	90% UCI
		PY 6 episodes (N)	Baseline mean	PY 6 mean	PY 6 episodes (N)	Baseline mean	PY 6 mean				
Mortality rate	Patients who were dually eligible	3,111	0.8%	0.7%	2,586	0.8%	0.8%	-0.1 pp	0.80	-0.5 pp	0.4 pp
	Patients who were not dually eligible	44,013	0.5%	0.4%	48,629	0.5%	0.5%	0.0 pp	0.68	-0.1 pp	0.1 pp
ED use	Patients who were dually eligible	3,101	19.1%	17.3%	2,575	19.1%	15.2%	2.1 pp	0.14	-0.2 pp	4.4 pp
	Patients who were not dually eligible	43,898	11.4%	11.3%	48,496	10.8%	10.4%	0.2 pp	0.53	-0.4 pp	0.9 pp
Readmission rate	Patients who were dually eligible	3,097	11.0%	8.0%	2,571	10.4%	8.0%	-0.5 pp	0.39	-1.5 pp	0.5 pp
	Patients who were not dually eligible	43,875	7.7%	5.2%	48,475	7.3%	5.3%	-0.5 pp	0.14	-1.0 pp	0.1 pp

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 ending between October 1, 2021, and December 31, 2022 (PY 6 intervention).

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

CJR = Comprehensive Care for Joint Replacement; DiD = difference-in-differences; DDD = difference-in-difference-in-differences; ED = emergency department; LCI = lower confidence interval; PY = performance year; UCI = upper confidence interval.

Exhibit F-6B: The Model Did Not Have Any Significant Impacts on Quality Measures for Patients Who Were or Were Not Dually Eligible, and Differences Between the Populations Were Not Statistically Significant

Outcome	Impact on:		Difference in impact	p-value	90% LCI	90% UCI
	Patients who were dually eligible	Patients who were not dually eligible				
Mortality rate	-0.1 pp	0.0 pp	0.0 pp	0.88	-0.5 pp	0.4 pp
ED use	2.1 pp	0.2 pp	1.8 pp	0.16	-0.3 pp	4.0 pp
Readmission rate	-0.5 pp	-0.5 pp	-0.1 pp	0.91	-1.1 pp	1.0 pp

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 ending between October 1, 2021, and December 31, 2022 (PY 6 intervention).

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

CJR = Comprehensive Care for Joint Replacement; DDD = difference-in-difference-in-differences; ED = emergency department; LCI = lower confidence interval; PY = performance year; UCI = upper confidence interval.

Exhibit F-7A: Patients Who Were Dually Eligible Had a Statistically Significant Increase in First Discharge With Home Health, but No Statistically Significant Differences Existed Relative to Patients Who Were Not Dually Eligible

First PAC discharge	Population	CJR			Control			Impact (DiD)	p-value	90% LCI	90% UCI
		PY 6 episodes (N)	Baseline mean	PY 6 mean	PY 6 episodes (N)	Baseline mean	PY 6 mean				
IRF	Patients who were dually eligible	3,103	13.8%	4.3%	2,577	12.2%	6.9%	-4.1 pp	0.13	-8.6 pp	0.4 pp
	Patients who were not dually eligible	43,908	10.9%	1.9%	48,502	10.1%	4.6%	-3.5 pp	0.06	-6.6 pp	-0.4 pp
SNF	Patients who were dually eligible	3,103	49.4%	19.0%	2,577	43.4%	17.7%	-4.8 pp	0.13	-10.0 pp	0.4 pp
	Patients who were not dually eligible	43,908	36.5%	7.6%	48,502	34.4%	8.8%	-3.4 pp	0.15	-7.2 pp	0.4 pp
HH	Patients who were dually eligible	3,103	29.7%	65.4%	2,577	32.4%	57.9%	10.2 pp	0.05	1.7 pp	18.6 pp
	Patients who were not dually eligible	43,908	38.7%	72.7%	48,502	35.3%	57.9%	11.2 pp*	0.09	0.5 pp	22.0 pp
No HH	Patients who were dually eligible	3,103	7.1%	11.3%	2,577	12.1%	17.5%	-1.2 pp	0.64	-5.5 pp	3.1 pp
	Patients who were not dually eligible	43,908	13.8%	17.8%	48,502	20.2%	28.6%	-4.4 pp	0.35	-12.2 pp	3.4 pp

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 ending between October 1, 2021, and December 31, 2022 (PY 6 intervention).

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

CJR = Comprehensive Care for Joint Replacement; DiD = difference-in-differences; DDD = difference-in-difference-in-differences; HH = home health; IRF = inpatient rehabilitation facility; LCI = lower confidence interval; pp = percentage point; PY = performance year; SNF = skilled nursing facility; UCI = upper confidence interval.

* The sample for this estimate failed tests for parallel trends in the baseline period.

Exhibit F-7B: Patients Who Were Dually Eligible Had a Statistically Significant Increase in First Discharge With Home Health, but No Statistically Significant Differences Existed Relative to Patients Who Were Not Dually Eligible

First PAC discharge	Impact on:		Difference in impact	p-value	90% LCI	90% UCI
	Patients who were dually eligible	Patients who were not dually eligible				
IRF	-4.1 pp	-3.5 pp	-0.6 pp	0.69	-3.2 pp	1.9 pp
SNF	-4.8 pp	-3.4 pp	-1.4 pp	0.46	-4.7 pp	1.8 pp
HH	10.2 pp	11.2 pp*	-1.1 pp	0.76	-6.9 pp	4.7 pp
No HH	-1.2 pp	-4.4 pp	3.2 pp	0.27	-1.6 pp	7.9 pp

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 ending between October 1, 2021, and December 31, 2022 (PY 6 intervention).

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

CJR = Comprehensive Care for Joint Replacement; DDD = difference-in-difference-in-differences; HH = home health; IRF = inpatient rehabilitation facility; LCI = lower confidence interval; PY = performance year; SNF = skilled nursing facility; UCI = upper confidence interval.

* The sample for this estimate failed tests for parallel trends in the baseline period.

Exhibit F-8A: Patients Who Were Dually Eligible Had a Statistically Significant Decrease in SNF Length of Stay, and the Difference Relative to Patients Who Were Not Dually Eligible Was Statistically Significant

Outcome	Population	CJR			Control			Impact (DiD)	p-value	90% LCI	90% UCI
		PY 6 episodes (N)	Baseline mean	PY 6 mean	PY 6 Episodes (N)	Baseline mean	PY 6 mean				
HH visits	Patients who were dually eligible	2,245	19.5	12.2	1,673	18.8	11.9	-0.5*	0.58	-1.9	0.9
	Patients who were not dually eligible	30,902	17.4	9.7	27,446	16.8	9.8	-0.7*	0.24	-1.8	0.3
SNF LOS	Patients who were dually eligible	714	26.9	24.7	471	25.0	25.8	-2.8*	0.02	-4.9	-0.8
	Patients who were not dually eligible	3,899	19.0	15.5	3,910	19.5	16.6	-0.6	0.49	-2.0	0.8
IRF LOS	Patients who were dually eligible	179	10.6	11.3	174	10.7	11.5	-0.2	0.81	-1.5	1.1
	Patients who were not dually eligible	1,172	10.1	11.4	1,806	10.1	10.8	0.5	0.07	0.0	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 ending between October 1, 2021, and December 31, 2022 (PY 6 intervention).

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

CJR = Comprehensive Care for Joint Replacement; DiD = difference-in-differences; DDD = difference-in-difference-in-differences; HH = home health; IRF = inpatient rehabilitation facility; LCI = lower confidence interval; LOS = length of stay; PY = performance year; SNF = skilled nursing facility; UCI = upper confidence interval.

* The sample for this estimate failed tests for parallel trends in the baseline period.

Exhibit F-8B: Patients Who Were Dually Eligible Had a Statistically Significant Decrease in SNF Length of Stay, and the Difference Relative to Patients Who Were Not Dually Eligible Was Statistically Significant

Outcome	Impact on:		Difference in impact	p-value	90% LCI	90% UCI
	Patients who were dually eligible	Patients who were not dually eligible				
HH visits	-0.5*	-0.7*	0.3	0.78	-1.3	1.8
SNF LOS	-2.8*	-0.6	-2.2	0.08	-4.3	-0.2
IRF LOS	-0.2	0.5	-0.7	0.36	-1.9	0.6

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 ending between October 1, 2021, and December 31, 2022 (PY 6 intervention).

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

CJR = Comprehensive Care for Joint Replacement; DDD = difference-in-difference-in-differences; HH = home health; IRF = inpatient rehabilitation facility; LCI = lower confidence interval; LOS = length of stay; PY = performance year; SNF = skilled nursing facility; UCI = upper confidence interval.

* The sample for this estimate failed tests for parallel trends in the baseline period.

Exhibit F-9A: Patients Who Are Black or African American and Dually Eligible Had No Significant Impacts for ED Use or Unplanned Readmission Rate, and No Significant Differences Relative to Patients Who Are Non-Hispanic White and Not Dually Eligible

Outcome	Population	CJR			Control			Impact (DiD)	p-value	90% LCI	90% UCI
		PY 6 episodes (N)	Baseline mean	PY 6 mean	PY 6 episodes (N)	Baseline mean	PY 6 mean				
Mortality rate	Patients who are Black or African American and dually eligible	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A
	Patients who are Non-Hispanic White and not dually eligible	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A
ED use	Patients who are Black or African American and dually eligible	396	24.0%	21.8%	483	23.9%	17.2%	4.5 pp	0.14	-0.5 pp	9.4 pp
	Patients who are Non-Hispanic White and not dually eligible	38,587	11.3%	11.3%	43,393	10.7%	10.7%	0.0 pp	0.94	-0.7 pp	0.7 pp
Readmission rate	Patients who are Black or African American and dually eligible	395	12.7%	8.8%	481	11.2%	7.7%	-0.4 pp	0.84	-3.3 pp	2.6 pp
	Patients who are Non-Hispanic White and not dually eligible	38,568	7.6%	5.3%	43,373	7.3%	5.4%	-0.5 pp	0.12	-1.0 pp	0.0 pp

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 ending between October 1, 2021, and December 31, 2022 (PY 6 intervention).

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

CJR = Comprehensive Care for Joint Replacement; DiD = difference-in-differences; DDD = difference-in-difference-in-differences; ED = emergency department; LCI = lower confidence interval; PY = performance year; UCI = upper confidence interval.

Exhibit F-9B: Patients Who Are Black or African American and Dually Eligible Had No Significant Impacts for ED Use or Unplanned Readmission Rate, and No Significant Differences Relative to Patients Who Are Non-Hispanic White and Not Dually Eligible

Outcome	Impact on:		Difference in impact	p-value	90% LCI	90% UCI
	Patients who are Black or African American and dually eligible	Patients who are Non-Hispanic White and not dually eligible				
Mortality rate	N/A	N/A	N/A	N/A	N/A	N/A
ED use	4.5 pp	0.0 pp	4.4 pp	0.16	-0.7 pp	9.6 pp
Readmission rate	-0.4 pp	-0.5 pp	0.1 pp	0.95	-2.7 pp	2.9 pp

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 ending between October 1, 2021, and December 31, 2022 (PY 6 intervention).

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

CJR = Comprehensive Care for Joint Replacement; DDD = difference-in-difference-in-differences; ED = emergency department; LCI = lower confidence interval; PY = performance year; UCI = upper confidence interval.

Exhibit F-10A: Patients Who Are Black or African American and Dually Eligible Had No Statistically Significant Impacts for First Discharge Destination, but the Differential Impact for Discharge Without Home Health Was Significant

First PAC discharge	Population	CJR			Control			Impact (DiD)	p-value	90% LCI	90% UCI
		PY 6 episodes (N)	Baseline mean	PY 6 mean	PY 6 episodes (N)	Baseline mean	PY 6 mean				
IRF	Patients who are Black or African American and dually eligible	396	16.5%	6.3%	483	12.8%	6.8%	-4.3 pp	0.15	-9.3 pp	0.6 pp
	Patients who are Non-Hispanic White and not dually eligible	38,597	10.2%	2.0%	43,398	9.8%	4.9%	-3.3 pp	0.08	-6.4 pp	-0.2 pp
SNF	Patients who are Black or African American and dually eligible	396	41.7%	15.9%	483	37.7%	13.3%	-1.4 pp	0.79	-10.0 pp	7.3 pp
	Patients who are Non-Hispanic White and not dually eligible	38,597	37.3%	7.1%	43,398	34.8%	8.1%	-3.5 pp	0.14	-7.3 pp	0.4 pp
HH	Patients who are Black or African American and dually eligible	396	34.6%	62.8%	483	34.3%	58.6%	3.9 pp	0.65	-10.3 pp	18.1 pp
	Patients who are Non-Hispanic White and not dually eligible	38,597	38.9%	72.8%	43,398	35.7%	57.7%	11.8 pp*	0.07	1.0 pp	22.7 pp
No HH	Patients who are Black or African American and dually eligible	396	7.2%	15.0%	483	15.3%	21.3%	1.8 pp	0.73	-6.8 pp	10.5 pp
	Patients who are Non-Hispanic White and not dually eligible	38,597	13.6%	18.1%	43,398	19.7%	29.3%	-5.1 pp	0.29	-13.0 pp	2.9 pp

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 ending between October 1, 2021, and December 31, 2022 (PY 6 intervention).

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

CJR = Comprehensive Care for Joint Replacement; DiD = difference-in-differences; DDD = difference-in-difference-in-differences; HH = home health; IRF = inpatient rehabilitation facility; LCI = lower confidence interval; PY = performance year; SNF = skilled nursing facility; UCI = upper confidence interval.

* The sample for this estimate failed tests for parallel trends in the baseline period.

Exhibit F-10B: Patients Who Are Black or African American and Dually Eligible Had No Statistically Significant Impacts for First Discharge Destination but Did Have a Statistically Significant Difference in Impacts for First Discharge to Home Without Home Health

First PAC discharge	Impact on:		Difference in impact	p-value	90% LCI	90% UCI
	Patients who are Black or African American and dually eligible	Patients who are Non-Hispanic White and not dually eligible				
IRF	-4.3 pp	-3.3 pp	-1.0 pp*	0.63	-4.5 pp	2.5 pp
SNF	-1.4 pp	-3.5 pp	2.1 pp	0.62	-4.9 pp	9.1 pp
HH	3.9 pp	11.8 pp*	-8.0 pp	0.14	-16.9 pp	1.0 pp
No HH	1.8 pp	-5.1 pp	6.9 pp	0.02	1.9 pp	11.9 pp

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 ending between October 1, 2021, and December 31, 2022 (PY 6 intervention).

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

CJR = Comprehensive Care for Joint Replacement; DDD = difference-in-difference-in-differences; HH = home health; IRF = inpatient rehabilitation facility; LCI = lower confidence interval; PY = performance year; SNF = skilled nursing facility; UCI = upper confidence interval.

* The sample for this estimate failed tests for parallel trends in the baseline period.

Exhibit F-11A: Patients Who Are Black or African American and Dually Eligible Had No Statistically Significant Impacts for PAC Utilization Measures and No Significant Differences Compared With Patients Who Are Non-Hispanic White and Not Dually Eligible

Outcome	Population	CJR			Control			Impact (DiD)	p-value	90% LCI	90% UCI
		PY 6 episodes (N)	Baseline mean	PY 6 mean	PY 6 episodes (N)	Baseline mean	PY 6 mean				
HH visits	Patients who are Black or African American and dually eligible	278	19.3	11.2	333	18.3	11.3	-1.0	0.30	-2.7	0.6
	Patients who are Non-Hispanic White and not dually eligible	27,234	17.3	9.6	24,534	16.6	9.7	-0.8*	0.21	-1.8	0.3
SNF LOS	Patients who are Black or African American and dually eligible	78	26.5	23.5	74	24.1	25.8	-4.6*	0.20	-10.6	1.4
	Patients who are Non-Hispanic White and not dually eligible	3,423	19.0	13.9	3,524	19.4	14.8	-0.5	0.52	-1.9	0.8
IRF LOS	Patients who are Black or African American and dually eligible	30	10.5	10.5	31	10.6	12.0	-1.4	0.29	-3.7	0.8
	Patients who are Non-Hispanic White and not dually eligible	1,010	10.2	11.0	1,583	10.1	10.5	0.4	0.16	-0.1	0.8

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 ending between October 1, 2021, and December 31, 2022 (PY 6 intervention).

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

CJR = Comprehensive Care for Joint Replacement; DiD = difference-in-differences; DDD = difference-in-difference-in-differences; HH = home health; IRF = inpatient rehabilitation facility; LCI = lower confidence interval; LOS = length of stay; PY = performance year; SNF = skilled nursing facility; UCI = upper confidence interval.

* The sample for this estimate failed tests for parallel trends in the baseline period.

Exhibit F-11B: Patients Who Are Black or African American and Dually Eligible Had No Statistically Significant Impacts for PAC Utilization Measures, and No Significant Differences Compared With Patients Who Are Non-Hispanic White and Not Dually Eligible

Outcome	Impact on:		Difference in impact	p-value	90% LCI	90% UCI
	Patients who are Black or African American and dually eligible	Patients who are Non-Hispanic White and not dually eligible				
HH visits	-1.0	-0.8*	-0.3	0.83	-2.4	1.8
SNF LOS	-4.6	-0.5	-4.1	0.26	-10.2	1.9
IRF LOS	-1.4	0.4	-1.8	0.19	-4.1	0.4

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 ending between October 1, 2021, and December 31, 2022 (PY 6 intervention).

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

CJR = Comprehensive Care for Joint Replacement; DDD = difference-in-difference-in-differences; HH = home health; IRF = inpatient rehabilitation facility; LCI = lower confidence interval; LOS = length of stay; PY = performance year; SNF = skilled nursing facility; UCI = upper confidence interval.

* The sample for this estimate failed tests for parallel trends in the baseline period.

Exhibit F-12A: Patients Who Are Hispanic Had a Significant Increase in ED Use, and This Was Significantly Different Relative to Patients Who Are Non-Hispanic White. They Also Had an Estimated Increase in Readmission Rate, Which Was Significantly Different Relative to Patients Who Are Non-Hispanic White

Outcome	Population	CJR			Control			Impact (DiD)	p-value	90% LCI	90% UCI
		PY 6 episodes (N)	Baseline mean	PY 6 mean	PY 6 episodes (N)	Baseline mean	PY 6 mean				
Mortality rate	Patients who are Hispanic	2,312	0.6%	0.7%	1,330	0.3%	0.1%	0.3 pp	0.29	-0.1 pp	0.7 pp
	Patients who are Non-Hispanic White	40,160	0.5%	0.5%	45,030	0.5%	0.5%	0.0 pp	0.76	-0.1 pp	0.1 pp
ED use	Patients who are Hispanic	2,312	12.2%	14.6%	1,328	13.8%	13.0%	3.1 pp	<0.01	1.7 pp	4.5 pp
	Patients who are Non-Hispanic White	40,035	11.9%	11.6%	44,902	11.2%	10.9%	0.1 pp	0.90	-0.7 pp	0.8 pp
Readmission rate	Patients who are Hispanic	2,308	7.7%	5.9%	1,327	7.5%	4.3%	1.4 pp	0.11	0.0 pp	2.8 pp
	Patients who are Non-Hispanic White	40,016	8.0%	5.4%	44,881	7.6%	5.6%	-0.5 pp	0.08	-1.0 pp	0.0 pp

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 ending between October 1, 2021, and December 31, 2022 (PY 6 intervention).

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

CJR = Comprehensive Care for Joint Replacement; DiD = difference-in-differences; DDD = difference-in-difference-in-differences; ED = emergency department; LCI = lower confidence interval; PY = performance year; UCI = upper confidence interval.

Exhibit F-12B: Patients Who Are Hispanic Had a Strongly Significant Increase in ED Use, and This Was Significantly Different Relative to Patients Who Are Non-Hispanic White. They Also Had an Estimated Increase in Readmission Rate, Which Was Significantly Different Relative to Patients Who Are Non-Hispanic White

Outcome	Impact on:		Difference in impact	p-value	90% LCI	90% UCI
	Patients who are Hispanic	Patients who are Non-Hispanic White				
Mortality	0.3 pp	0.0 pp	0.3 pp	0.28	-0.1 pp	0.7 pp
ED use	3.1 pp	0.1 pp	3.0 pp	<0.01	1.5 pp	4.5 pp
Unplanned readmissions	1.4 pp	-0.5 pp	1.9 pp	0.03	0.5 pp	3.4 pp

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 ending between October 1, 2021, and December 31, 2022 (PY 6 intervention).

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

CJR = Comprehensive Care for Joint Replacement; DDD = difference-in-difference-in-differences; ED = emergency department; LCI = Lower Confidence Interval; PY = performance year; UCI = upper confidence interval.

Exhibit F-13A: Patients Who Are Hispanic Had a Significant Increase in First Discharge With Home Health and a Significant Decrease in First Discharge to SNF. Compared With Patients Who Are Non-Hispanic White, the Difference in Impact for First Discharge to SNF Was Significant

First PAC discharge	Population	CJR			Control			Impact (DiD)	p-value	90% LCI	90% UCI
		PY 6 episodes (N)	Baseline mean	PY 6 mean	PY 6 episodes (N)	Baseline mean	PY 6 mean				
IRF	Patients who are Hispanic	2,312	14.0%	2.9%	1,328	9.5%	4.2%	-5.8 pp	0.11	-11.7 pp	0.2 pp
	Patients who are Non-Hispanic White	40,046	10.7%	1.9%	44,909	10.3%	4.7%	-3.1 pp	0.10	-6.2 pp	-0.0 pp
SNF	Patients who are Hispanic	2,312	44.5%	8.7%	1,328	37.3%	9.8%	-8.3 pp	0.02	-14.2 pp	-2.3 pp
	Patients who are Non-Hispanic White	40,046	37.8%	7.7%	44,909	35.4%	8.8%	-3.5 pp	0.14	-7.4 pp	0.4 pp
HH	Patients who are Hispanic	2,312	32.5%	75.1%	1,328	38.0%	67.0%	13.6 pp	0.01	4.6 pp	22.6 pp
	Patients who are Non-Hispanic White	40,046	38.1%	73.0%	44,909	34.9%	58.6%	11.3 pp*	0.08	0.8 pp	21.9 pp
No HH	Patients who are Hispanic	2,312	9.0%	13.4%	1,328	15.1%	19.1%	0.5 pp	0.88	-4.5 pp	5.4 pp
	Patients who are Non-Hispanic White	40,046	13.5%	17.4%	44,909	19.4%	28.0%	-4.7 pp	0.30	-12.2 pp	2.8 pp

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 ending between October 1, 2021, and December 31, 2022 (PY 6 intervention).

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

CJR = Comprehensive Care for Joint Replacement; DiD = difference-in-differences; DDD = difference-in-difference-in-differences; HH = home health; IRF = inpatient rehabilitation facility; LCI = lower confidence interval; PY = performance year; SNF = skilled nursing facility; UCI = upper confidence interval.

* The sample for this estimate failed tests for parallel trends in the baseline period.

Exhibit F-13B: Patients Who Are Hispanic Had a Significant Increase in First Discharge With Home Health and a Significant Decrease in First Discharge to SNF. Compared With Patients Who Are Non-Hispanic White, the Difference in Impact for First Discharge to SNF Was Significant

First PAC discharge	Impact on:		Difference in impact	p-value	90% LCI	90% UCI
	Patients who are Hispanic	Patients who are Non-Hispanic White				
IRF	-5.8 pp	-3.1 pp	-2.6 pp	0.47	-8.7 pp	3.4 pp
SNF	-8.3 pp	-3.5 pp	-4.8 pp	0.10	-9.6 pp	-0.0 pp
HH	13.6 pp	11.3 pp*	2.3 pp	0.53	-3.7 pp	8.2 pp
No HH	0.5 pp	-4.7 pp	5.2 pp	0.11	-0.1 pp	10.4 pp

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 ending between October 1, 2021, and December 31, 2022 (PY 6 intervention).

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

CJR = Comprehensive Care for Joint Replacement; DDD = difference-in-difference-in-differences; HH = home health; IRF = inpatient rehabilitation facility; LCI = lower confidence interval; PY = performance year; SNF = skilled nursing facility; UCI = upper confidence interval.

* The sample for this estimate failed tests for parallel trends in the baseline period.

Exhibit F-14A: Patients Who Are Hispanic Had No Statistically Significant Impacts for Utilization Measures, nor Was Any Difference Statistically Significant Relative to Patients Who Are Non-Hispanic White

Outcome	Population	CJR			Control			Impact (DiD)	p-value	90% LCI	90% UCI
		PY 6 episodes (N)	Baseline mean	PY 6 mean	PY 6 episodes (N)	Baseline mean	PY 6 mean				
HH visits	Patients who are Hispanic	1,655	20.2	11.4	826	18.5	10.4	-0.7	0.51	-2.6	1.1
	Patients who are Non-Hispanic White	28,260	17.4	9.6	25,480	16.8	9.7	-0.7*	0.23	-1.7	0.3
SNF LOS	Patients who are Hispanic	237	23.6	20.7	109	22.5	21.3	-1.7	0.34	-4.6	1.3
	Patients who are Non-Hispanic White	3,799	19.6	16.0	3,852	20.0	16.9	-0.5	0.56	-1.9	0.9
IRF LOS	Patients who are Hispanic	95	10.3	11.5	55	10.9	11.6	0.4	0.57	-0.8	1.7
	Patients who are Non-Hispanic White	2,312	10.2	11.1	1,328	10.1	10.6	0.4	0.16	-0.1	0.8

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 ending between October 1, 2021, and December 31, 2022 (PY 6 intervention).

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

CJR = Comprehensive Care for Joint Replacement; DiD = difference-in-differences; DDD = difference-in-difference-in-differences; HH = home health; IRF = inpatient rehabilitation facility; LCI = lower confidence interval; LOS = length of stay; PY = performance year; SNF = skilled nursing facility; UCI = upper confidence interval.

* The sample for this estimate failed tests for parallel trends in the baseline period.

Exhibit F-14B: Patients Who Are Hispanic Have No Statistically Significant Impacts for Utilization Measures, nor Was Any Difference Statistically Significant Relative to Patients Who Are Non-Hispanic White

Outcome	Impact on:		Difference in impact	p-value	90% LCI	90% UCI
	Patients who are Hispanic	Patients who are Non-Hispanic White				
HH visits	-0.7	-0.7*	0.0	0.97	-1.6	1.5
SNF LOS	-1.7	-0.5	-1.2	0.44	-3.8	1.4
IRF LOS	0.4	0.4	0.0	0.95	-1.1	1.2

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 ending between October 1, 2021, and December 31, 2022 (PY 6 intervention).

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

CJR = Comprehensive Care for Joint Replacement; DDD = difference-in-difference-in-differences; HH = home health; IRF = inpatient rehabilitation facility; LCI = lower confidence interval; LOS = length of stay; PY = performance year; SNF = skilled nursing facility; UCI = upper confidence interval.

* The sample for this estimate failed tests for parallel trends in the baseline period.

B. Distribution of Underserved Populations Across Hospitals and Time in the CJR Model

In the CJR Model evaluation's Fifth Annual Report, we observed differential decreases in LEJR volume for some underserved populations, which motivated an investigation into hospitals that had large changes in the share of underserved populations.[122] In this section, we present descriptive statistics for CJR hospitals that had the largest changes in the share of patients from underserved populations between baseline period and PYs 5–6.

1. Summary of Findings

- Among CJR participants, hospitals with large increases in patient share from three of the four underserved populations (Black or African American, Black or African American and dually eligible, and Hispanic) had lower average baseline shares than hospitals with large decreases in the same patient share.
- Among control hospitals, those with large increases in patient share from all four underserved populations had lower average baseline patient share from those populations compared to control hospitals with large decreases in patient share from the same population.
- For both CJR and control hospitals, hospitals with large increases or decreases in the share of patients who are Black or African American were noticeably different from hospitals with large increases or decreases in the share of patients who were Hispanic and from hospitals with large increases or decreases in the share of patients who were dually eligible.

2. Methods

We analyzed the change in hospitals' share of patients from four underserved populations between the baseline period and PYs 5–6: patients who are Black or African American, patients who are dually eligible, patients who are Black or African American and dually eligible, and patients who are Hispanic.

CJR and control hospitals were ordered separately based on their change in share between the baseline and PY for each of the four underserved populations. Then, for each ordering, the 10th and 90th percentiles were calculated. Hospitals in the 90th percentile or above for a given underserved population were defined as hospitals with large increases in the underserved patient share. Hospitals in the 10th percentile or lower for a given underserved population were identified as hospitals with large decreases in the underserved patient share.

3. Results

In this analysis, we paid specific attention to hospitals that experienced large changes in their share of patients from underserved populations between the baseline period and PYs 5–6. Of particular interest was whether hospitals that experienced large increases (or large decreases) in their share of

[122] Centers for Medicare & Medicaid Services. (2023). *Comprehensive Care for Joint Replacement Model - Fifth Annual Report*. https://www.cms.gov/priorities/innovation/data-and-reports/2023/cjr-py5-annual-report

patients from a given underserved population were like hospitals that experienced similar changes in their share of patients from the other underserved populations: Was there a set of hospital characteristics that was correlated with large changes in the underserved population share for all four underserved populations?

Among CJR participants, hospitals with large increases in patient share from three of the four underserved populations (Black or African American, Black or African American and dually eligible, and Hispanic) had lower average baseline shares than hospitals with large decreases in the same patient share. Hospitals with large increases in patient share from two underserved populations (Black or African American and Black or African American and dually eligible) had lower PY 6 episode volume than hospitals with large decreases in the same patient share; hospitals with large increases in patient share from the other two populations (dually eligible and Hispanic) had similar or larger PY 6 episode volume than hospitals with large decreases. Hospitals with large increases in patient share from two of the four underserved populations (Black or African American and Hispanic) had lower average bed counts than hospitals with large decreases in the same share; hospitals with large increases in patient share from the other two populations (dually eligible and Black or African American and dually eligible) had higher average bed count **(Exhibit F-15, Exhibit F-16, Exhibit F-17, and Exhibit F-18).**

Exhibit F-15: The Gap Between Hospitals With Large Increases and Large Decreases in the Share of Patients Who Are Black or African American Shrank From Baseline to PYs 5–6

Descriptive	CJR		Control	
	Large positive	Large negative	Large positive	Large negative
Change in share of patients who are Black or African American (Baseline → PYs 5–6)[a]	8.6% → 16.1% +7.5 pp	25.7% → 12.5% -13.2 pp	12.4% → 21.0% +8.6 pp	34.2% → 19.0% -15.2 pp
PY 6 episode volume[a]	47.1	77.8	54.1	77.4
Bed count[a]	287	303	198	246
Regions	7 W, 12 S, 13 NE	6 W, 12 S, 6 NE, 4 MW	2 W, 22 S, 4 NE, 4 MW	1 W, 13 S, 5 NE, 13 MW

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), episodes ending between January 1, 2020, and September 30, 2021 (PY 5 intervention), and episodes ending between October 1, 2021, and December 31, 2022 (PY 6 intervention).

Notes: Hospitals with large positive changes in underserved patient share were defined as being in the 90th percentile or above for the change in a given underserved population share. Hospitals with large negative changes in underserved patient share were defined as being in the 10th percentile or lower for a given underserved population. Regions are U.S. census regions. MW = Midwest; NE = Northeast; S = South; W = West. CJR = Comprehensive Care for Joint Replacement; pp = percentage point; PY = performance year.

[a] Average per hospital.

Exhibit F-16: The Gap Between Hospitals With Large Positive and Large Negative Changes in the Share of Patients Who Were Dually Eligible Grew From Baseline to PYs 5–6

Descriptive	CJR		Control	
	Large positive	Large negative	Large positive	Large negative
Change in share of patients who were dually eligible (Baseline → PYs 5–6)[a]	37.8% → 51.4% +13.6 pp	44.2% → 21.4% -22.8 pp	16.9% → 24.4% +7.5 pp	35.2% → 17.5% -17.7 pp
PY 6 episode volume[a]	51.8	56.7	55.2	66.5
Bed count[a]	284	230	243	219
Regions	15 W, 7 S, 9 NE, 1 MW	8 W, 12 S, 8 NE, 4 MW	5 W, 10 S, 6 NE, 11 MW	2 W, 19 S, 3 NE, 8 MW

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), episodes ending between January 1, 2020, and September 30, 2021 (PY 5 intervention), and episodes ending between October 1, 2021, and December 31, 2022 (PY 6 intervention).

Notes: Hospitals with large positive changes in underserved patient share were defined as being in the 90th percentile or above for the change in a given underserved population share. Hospitals with large negative changes in underserved patient share were defined as being in the 10th percentile or lower for a given underserved population. Regions are U.S. census regions. MW = Midwest; NE = Northeast; S = South; W = West. CJR = Comprehensive Care for Joint Replacement; pp = percentage point; PY = performance year.

[a] Average per hospital.

Among control hospitals, those with large increases in patient share from all four underserved populations had lower average baseline patient share from those populations compared to control hospitals with large decreases in patient share from the same population. Hospitals with large increases in patient share from three of the four underserved populations (Black or African American, dually eligible, Hispanic) had lower PY 6 episode volume than hospitals with large decreases in patient share from the same population. Hospitals with large increases in patient share from all four populations had slightly lower average bed count compared to hospitals with large decreases in patient share from the same population, but the largest difference was much smaller than it was for CJR participant hospitals.

Exhibit F-17: Hospitals With High Positive and High Negative Changes in Share of Patients Who Are Both Black or African American and Dually Eligible

Descriptive	CJR		Control	
	Large positive	Large negative	Large positive	Large negative
Change in share of patients who are Black or African American and dually eligible (Baseline → PYs 5–6)[a]	4.4% → 8.5% +4.1 pp	17.6% → 7.0% -10.6 pp	3.6% → 7.8% +4.2 pp	19.5% → 8.3% -11.2 pp
PY 6 episode volume[a]	34.5	58.8	61.4	55.1
Bed count[a]	287	276	225	263
Regions	10 W, 12 S, 10 NE	5 W, 16 S, 9 NE, 2 MW	2 W, 17 S, 3 NE, 10 MW	20 S, 3 NE, 9 MW

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), episodes ending between January 1, 2020, and September 30, 2021 (PY 5 intervention), and episodes ending between October 1, 2021, and December 31, 2022 (PY 6 intervention).

Notes: Hospitals with large positive changes in underserved patient share were defined as being in the 90th percentile or above for the change in a given underserved population share. Hospitals with large negative changes in underserved patient share were defined as being in the 10th percentile or lower for a given underserved population. Regions are U.S. census regions. MW = Midwest; NE = Northeast; S = South; W = West. CJR = Comprehensive Care for Joint Replacement; pp = percentage point; PY = performance year.

[a] Average per hospital.

For both CJR and control hospitals, hospitals with large increases or decreases in the share of patients who are Black or African American were noticeably different from hospitals with large increases or decreases in the share of patients who were Hispanic and from hospitals with large increases or decreases in the share of patients who were dually eligible. Hospitals with large increases or decreases in the share of patients who are Black or African American and dually eligible were more like hospitals with large increases or decreases in the share of patients who were dually eligible than to hospitals with large increases or decreases in the share of patients who are Black or African American.

Exhibit F-18: Hospitals With High Positive and High Negative Changes in Share of Patients Who Are Hispanic

Descriptive	CJR		Control	
	Large positive	Large negative	Large positive	Large negative
Change in share of patients who are Hispanic (Baseline → PYs 5–6)[a]	13.6% → 20.0% +6.4 pp	42.0% → 26.9% -15.1 pp	8.3% → 13.4% +5.1 pp	14.3% → 6.8% -7.5 pp
PY 6 episode volume[a]	79.6	63.0	48.5	55.6
Bed count[a]	219	297	174	184
Regions	13 W, 11 S, 8 NE	17 W, 11 S, 4 NE	6 W, 16 S, 4 NE, 5 MW	4 W, 19 S, 2 NE, 7 MW

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), episodes ending between January 1, 2020, and September 30, 2021 (PY 5 intervention), and episodes ending between October 1, 2021, and December 31, 2022 (PY 6 intervention).

Notes: Hospitals with large positive changes in underserved patient share were defined as being in the 90th percentile or above for the change in a given underserved population. Hospitals with large negative changes in underserved patient share were defined as being in the 10th percentile or lower for a given underserved population. Regions are U.S. census regions. MW = Midwest; NE = Northeast; S = South; W = West. CJR = Comprehensive Care for Joint Replacement; pp = percentage point; PY = performance year.

[a] Average per hospital.