LewinGroup®

CMS Comprehensive Care for Joint Replacement Model: Performance Year 5 Evaluation Report – Appendices

Fifth Annual Report

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

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

April 2023

CMS Comprehensive Care for Joint Replacement Model: Performance Year 5 Evaluation Report – Appendices

Fifth Annual Report

List of Appendices

Appendix A: List of Acronyms & Glossary Terms

Exhibit A-1: List of acronyms

Acronym	Meaning
ACH	Acute Care Hospital
ACO	Accountable Care Organization
ADLs	Activities of Daily Living
APM	Alternative Payment Model
ASC	Ambulatory Surgical Center
BPCI	Bundled Payments for Care Improvement
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
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
LEJR	Lower Extremity Joint Replacement
LOS	Length of Stay
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
NPRA	Net Payment Reconciliation Amount

Acronym	Meaning
OP	Outpatient
OT	Occupational Therapy
PAC	Post-Acute Care
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
SNF	Skilled Nursing Facility
THA	Total Hip Arthroplasty
TKA	Total Knee Arthroplasty
VBP	Value-Based Payments

Exhibit A-2: Glossary of terms

Term	Definition
90-day post-discharge period (PDP)	The 90 days following discharge from the anchor hospitalization.
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	The hospitalization 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	A 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 PY2/3, the effective discount percentages were: 0.5% for "excellent" quality, 1% for "good" quality, and 2% for "acceptable" or "below acceptable" quality. For hospitals with repayment responsibility in PY4/5, the effective discount percentages are: 1.5% for "excellent" quality, 2% for "good" quality, and 3% for "acceptable" and "below acceptable" quality.
Episode benchmark price	The episode benchmark price represents the expected episode payments if treatment patterns and patient mix did not change from historical spending for LEJR episodes. In the first three years of the model, the episode benchmark price is based on a blend of hospital-specific and regional historical LEJR payments. In PY4/5, the episode benchmark price is based solely on regional amounts. The product of the episode benchmark price and the effective discount percentage equals the quality-adjusted target price.
Episode of care	For the CJR model, an episode of care is triggered by an inpatient hospitalization for an LEJR procedure in which a beneficiary is discharged 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.

Term	Definition
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)	Counties associated with a core urban area that 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. Under this waiver, CMS waives the direct supervision requirement for home visits so that CJR beneficiaries may receive a limited number of home visits (up to nine per episode) by licensed clinical staff paid under the Medicare Physician Fee Schedule.
Post-discharge period (PDP)	Period of time starting on the day of the anchor hospitalization discharge. For the CJR model, the post-discharge period covers the 90 days after discharge.
Post-episode care	Under the CJR model, care that occurs after the 90-day post-discharge period.
Quality-adjusted target price	The quality-adjusted target price is based on three years of historical data and is a blend of the hospital historical episode payments and the regional average historical payments in the first three years of the CJR model. In PY4/5, the target price is based completely on the regional historical episode payment. The three years of historical data is rolling across performance years (2012-2014 for years 1 and 2, 2014-2016 for years 3 and 4, 2016-2018 for year 5). The quality adjustment at the beginning of the performance year assumes that the hospital's composite quality score falls in the "acceptable" range. The quality adjustment reflects 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.
Reconciliation payment	A retrospective payment that Medicare makes to a CJR participant hospital if total fee-for-service payments for its episodes during a performance year are less than the aggregate quality-adjusted target price. 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 in PY2-5.
Related items and services	Episode-related items and services paid under Medicare Part A or Part B, after exclusions are applied, that are included in the bundle. These 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 take into account 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.

Term	Definition
Stop-loss/Stop-gain limits	Adjustments included in the NPRA calculation that 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 limits are 5% in PY2, 10% in PY3, and 20% in PY4 and PY5. The stop-gain limits are 5% in PY1 and PY2, 10% in PY3, and 20% in PY4 and PY5.
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: CJR Programmatic Flexibilities, Including Financial Arrangements, Beneficiary Incentives, and Program Rule Waivers

The CJR model allows hospitals to use fraud and abuse waivers issued by the Department of Health and Human Services to facilitate the implementation of care redesign interventions. Participating hospitals may or may not elect to use these waivers. Under the CJR model, hospitals may enter into financial arrangements with CJR collaborators, collaboration agents, downstream collaboration agents or provide incentives to CJR beneficiaries. Additionally, CMS waives certain Medicare program rules for beneficiaries in CJR episodes, such as: the direct supervision requirement for post-discharge home visits, specific requirements for furnishing telehealth services, and the three-day hospital stay requirement for coverage of skilled nursing facility (SNF) care. These waivers allow CJR beneficiaries to receive services under circumstances that would not otherwise be covered by Medicare.

The waivers allowed under the CJR model include:

- **Financial Arrangements –** Under the CJR model, hospitals may enter into sharing arrangements with certain collaborating providers and suppliers that are engaged in care redesign with the hospital and that furnish services to the beneficiary during an episode. Under such a sharing arrangement, hospitals may pass on a portion of their reconciliation payment, internal cost savings, or both (i.e., a gainsharing payment) to collaborating providers and suppliers. Sharing arrangements may also permit payments from a CJR collaborator to a participant hospital (i.e., an alignment payment) when the participating hospital has to repay CMS. Collaborators may be a SNF, home health agency (HHA), long-term care hospital (LTCH), inpatient rehabilitation facility (IRF), comprehensive outpatient rehabilitation facility, therapist in private practice, physician, non-physician practitioner, provider or supplier of outpatient therapy services, physician group practice (PGP), non-physician provider group practice, therapy group practice, accountable care organization (ACO), hospital, or critical access hospital. Under the CJR model, gainsharing payments must be made according to a pre-specified methodology.

  To be eligible to receive a gainsharing payment, collaborators must meet quality criteria for the performance year for which the participant hospital accrued the internal cost savings or earned the reconciliation payment that comprises the gainsharing payment. The quality of care criteria must be established by the participant hospital and directly related to the CJR episode. A CJR collaborator other than an ACO, PGP, non-physician provider group practice, or therapy group practice must have directly furnished a billable item or service to a CJR beneficiary during a CJR episode that occurred in the same performance year for which the participant hospital accrued the internal cost savings or earned the reconciliation payment that comprises the gainsharing payment or was assessed a repayment amount. A CJR collaborator that is a PGP, non-physician provider group practice, or therapy group practice must have billed for an item or service that was rendered by one or more PGP member, non-physician provider group practice member, or

therapy group practice member respectively to a CJR beneficiary during a CJR episode that occurred during the same performance year for which the participant hospital accrued the internal cost savings or earned the reconciliation payment that comprises the gainsharing payment or was assessed a repayment amount and must have contributed to CJR activities and been clinically involved in the care of CJR beneficiaries during the same performance year for which the CJR participant hospital accrued the internal cost savings or earned the reconciliation payment that comprises the gainsharing payment or was assessed a repayment amount. A CJR collaborator that is an ACO must have had an ACO provider/supplier that directly furnished, or an ACO participant that billed for, an item or service that was rendered to a CJR beneficiary during a CJR episode that occurred during the same performance year for which the participant hospital accrued the internal cost savings or earned the reconciliation payment that comprises the gainsharing payment or was assessed a repayment amount and the ACO must have contributed to CJR activities and been clinically involved in the care of CJR beneficiaries during the same performance year for which the participant hospital accrued the internal cost savings or earned the reconciliation payment that comprises the gainsharing payment or was assessed the repayment amount. In the event that a hospital is due to make a repayment to CMS under the CJR model, the total amount of alignment payments received by the hospital from a CJR collaborator that is an ACO may not be greater than 50% of the amount the hospital owes CMS. With respect to a CJR collaborator other than an ACO, the total amount of alignment payments received by the hospital may not be greater than 25% percent of the amount the hospital owes CMS. CMS also requires that gainsharing agreements cannot incentivize CJR collaborators to reduce service or provide substandard care to Medicare beneficiaries.

- **Beneficiary Incentives –** Participating hospitals may provide certain in-kind items or services to CJR beneficiaries during an episode of care. The item or service must be reasonably connected to a beneficiary's medical care and either be preventive or advance a clinical goal. Incentives may include technology items, which can be used for telehealth visits.

- **Post-Discharge Home Visit Waiver –** The direct supervision requirement for home visits can be waived so that CJR beneficiaries may receive a limited number of home visits (up to nine post-discharge home visits per episode) by licensed clinical staff paid under the Medicare Physician Fee Schedule.

- **Telehealth Waiver –** Under the CJR model, geographic and originating site requirements that typically apply for Medicare coverage of telehealth services may be waived as long as services are furnished according to other coverage and payment criteria. Medicare coverage criteria typically require telehealth services be furnished to individuals in certain geographic areas, including rural, medically underserved areas. For the CJR model, CMS waived this provision, allowing Medicare coverage of telehealth services furnished to eligible beneficiaries regardless of their geographic region. Medicare coverage criteria also specify that Medicare may only cover telehealth services that are received in certain

clinical settings. For the CJR model, the originating site requirement is waived for eligible beneficiaries receiving telehealth services from their homes or places of residence.

During the COVID-19 Public Health Emergency (PHE), CMS allowed Medicare coverage of telehealth services furnished to all patients regardless of their geographic locale and covered services, including physical and occupational therapy.[1] This waiver provided control hospitals with the same flexibility to utilize telehealth and ends with the conclusion of the PHE (which is ongoing).

- **Waiver of SNF 3-Day Rule –** Under traditional Medicare fee-for-service (FFS) rules, beneficiaries are not eligible for Medicare-covered SNF care unless they have a prior inpatient hospital stay of at least three consecutive days within 30 days of SNF admission. Under the SNF 3-day waiver, CJR participant hospitals can discharge a CJR beneficiary to an approved SNF without a qualifying 3-day inpatient stay when medically appropriate. This waiver became available in performance year 2 of the CJR model. A provision of this waiver is CJR beneficiaries may only be discharged to a SNF that is approved at the time of the beneficiary's admission. An approved SNF is one that received three or more stars on CMS' Five-Star Quality Rating System[2] for at least seven out of the past twelve months. CMS maintains a list of approved SNFs based on these requirements on the CJR model web site, which is updated quarterly.[3]

During the COVID-19 PHE, CMS waived the requirement for a 3-day prior hospitalization for coverage of a SNF stay for all hospitals in the United States, not just CJR participant hospitals.[1] Unlike the CJR model waiver, there was no requirement that the SNF had to meet minimum quality standards. This waiver provided control hospitals with the same flexibility to discharge patients with less than a 3-day hospital stay to a SNF and ends with the conclusion of the PHE (which is ongoing).

---

[1] COVID-19 Emergency Declaration Blanket Waivers for Health Care Providers (cms.gov)
[2] www.medicare.gov/NursingHomeCompare/
[3] https://innovation.cms.gov/innovation-models/cjr

Appendix C: Methodology

Table of Contents

I.   Data Sources

A.   Secondary data sources

Secondary data sources were used to:

1) Identify and characterize CJR participant hospitals and control group hospitals for risk adjustment and creation of weights for mandatory CJR hospitals and matched control groups for hospitals in voluntary metropolitan statistical areas (MSA) (Provider of Services file, Acute IPPS Final Rule data files, Medicare FFS claims, CJR programmatic data, Bundled Payments for Care Improvement Salesforce Database, Bundled Payments for Care Improvement Advanced Participant Database, AHRQ Compendium of U.S. Health Systems, and USAFacts Coronavirus Live Map);

2) Sample CJR participant hospitals for participation in telephone interviews and surveys (CJR programmatic data and Medicare FFS claims);

3) Identify lower extremity joint replacement (LEJR) discharges, create LEJR episodes, characterize episodes and beneficiaries, and evaluate the impact of the CJR model on health equity and LEJR discharge volume (Medicare FFS claims, Medicare FFS beneficiary enrollment data, Master Data Management (MDM), Bundled Payments for Care Improvement Salesforce Database, and Bundled Payments for Care Improvement Advanced Participant Database); and

4) Generate payment, utilization, quality, and savings to Medicare (Medicare FFS claims, Medicare standardized payments, and CJR programmatic data).

Exhibit C-1 lists the secondary sources, their contents, purpose in this evaluation, and relevant date ranges used for this report.

Exhibit C-1:   Secondary data sources

Data source	Date range	Dataset contents	Use
Area Health Resource Files (AHRF)	2015-2016 (Data is from 2012-2014)	County-level data aggregated to the MSA level. Variables include Medicare Advantage penetration, average Medicare beneficiary hierarchical condition category (HCC) score, dual eligible percentage, population per square mile, geography, and supply of health care facilities (SNF beds, LTCH beds) and health care professionals (primary care physicians, orthopedic surgeons, NPs/PAs, specialists).	Used to control for MSA Medicare Advantage penetration in the patient survey analysis.

Data source	Date range	Dataset contents	Use
Agency for Healthcare Research and Quality (AHRQ) Compendium of U.S. Health Systems	2016	Includes information on U.S. health systems, which are defined as having at least one hospital and at least one group of physicians providing comprehensive care and connected with each other through common ownership or joint management. Provides a crosswalk of hospital CCN to health system.	Used to characterize CJR and control hospitals and respondents of the care coordination survey.
Bundled Payments for Care Improvement Advanced Participant Database	Intervention	Identifies health care providers (hospitals, physicians, and physician practice groups) that are participating in the Bundled Payments for Care Improvement Advanced initiative, 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 Bundled Payments for Care Improvement Advanced participants for risk adjustment.
Bundled Payments for Care Improvement Participant Database	Baseline and intervention	Identifies health care providers (hospitals, PAC providers, physicians, and physician practice groups) that are participating in the Bundled Payments for Care Improvement 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 Bundled Payments for Care Improvement participants for exclusion. Used to identify hospitals as past Bundled Payments for Care Improvement LEJR participants for risk adjustment, creation of propensity score weights (PSW), and creation of matched control groups for hospitals in voluntary MSAs. Used to create a measure of Bundled Payments for Care Improvement dose for the volume analysis.
CJR programmatic data	Intervention	List of CJR participant hospitals, as well as their PY1, PY2, PY3 PY4, PY5.1, and PY5.2 quality-adjusted target prices, reconciliation (net payment reconciliation amount or NPRA), and hospital quality data.	Used to identify CJR participating hospitals, hospitals that continued mandatory participation in PY3, 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.
FY Acute IPPS Final Rule data files	FY 2016 (Data is from FY 2012-2014)	On an annual basis, CMS sets acute care hospital IPPS payment rates. Data files include fiscal year hospital-level information on provider identification number, bed count, medical residents per 1,000 beds, average daily census, DSH patient percentage, UCP per claim, Medicare days as a percent of total inpatient days, and section 401 status.	Used to risk adjust for acute care IPPS hospital characteristics. Used in the creation of PSW and matched control groups for hospitals in voluntary MSAs. Used to identify section 401 hospitals located in control group hospitals to exclude from the mandatory analysis.

Data source	Date range	Dataset contents	Use
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 Medicare Shared Savings 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, CEC with downside risk, and Next Generation).
Medicare FFS beneficiary enrollment data	Baseline and Intervention	Enrollment data (from CME and MBSF) provide beneficiary Medicare Part A/B eligibility information.	Enrollment data were used to confirm beneficiary eligibility and provide beneficiary characteristics for analyses (e.g., risk adjustment models, LEJR volume analysis, creation of PSW and matched control groups for hospitals in voluntary MSAs). 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	Parts A and B claims data (from TAP files) provide claims for different services received during the anchor hospitalization and post-discharge period (e.g., dates and types of service). A minimum three-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 (e.g., beneficiary prior utilization, HCC score, COVID-19 diagnosis) and outcome variables (e.g., unplanned readmissions, emergency department visits, and number of days/visits in each PAC setting); 2) create PSW for hospitals in mandatory MSAs and matched control groups for hospitals in voluntary MSAs; 3) generate the number of LEJR discharges at the market level; 4) 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 5) sample participants for primary data collection (patient survey, telephone interviews).

Data source	Date range	Dataset contents	Use
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 five, 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/emotional, bladder, bowel, swallowing/nutritional, and pain status. A minimum six 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 six months preceding the episode, evaluate nursing facility and SNF use for fracture episodes during the episodes and one-year after discharge, and measure the change in case-mix of CJR patients and patients in the control group who were discharged from the hospital to a nursing facility or SNF, between the baseline and the intervention periods.
Medicare standardized payments	Baseline and Intervention	Medicare standardized payments for 100% of Part A and B claims received via the IDR. Produced by a CMS contractor.	Used to create Medicare standardized paid amounts (Part 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.
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 creation of PSW and matched control groups for hospitals in voluntary MSAs.
USAFacts	2020 and later	Information on COVID-19 infection and mortality rates at the national, state and county-levels. Aggregated from individual federal, state and local data systems.	Used to create risk adjustment variables 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.

Note: ACO = accountable care organization, AHRF = Area Health Resource Files, AHRQ = Agency for Healthcare Research and Quality, CEC = comprehensive ESRD care model, CME = common Medicare enrollment, CMS = Centers for Medicare & Medicaid Services, COVID-19 = coronavirus disease 2019, DSH = disproportionate share hospital, ESRD = end-stage renal disease, FFS = fee-for-service, FY = fiscal year, HCC = hierarchical condition category, IDR = integrated data repository, IPPS = Inpatient Prospective Payment System, LEJR = lower extremity joint replacement, MBSF = Medicare beneficiary summary file, 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, SNF = skilled nursing facility, TAP = monthly Medicare claims file, THA = total hip arthroplasty, TKA = total knee arthroplasty, UCP = uncompensated care payment.

B. Primary data sources

To inform questions that are not readily answered by secondary data, we collected and analyzed primary data from telephone interviews with CJR participant hospitals and a survey of hospital care coordinators. We conducted one round of telephone interviews with 39 hospitals and received 199 survey responses from care coordinators at participating hospitals. In this appendix, we

describe the methods employed during the fifth performance year. Prior primary data collection efforts are detailed in prior annual reports.[1, 2, 3, 4]

1. Provider telephone interviews

In PY5, we conducted one round of telephone interviews with mandatory CJR participant hospitals. The aim of the interviews was to explore how the CJR model affected hospitals' experiences caring for beneficiaries with LEJR due to hip fracture.

a. Interviewees

The team interviewed representatives from current CJR participant hospitals (as of November 2021) that cared for patients who had LEJR due to hip fracture.

b. Protocols

We developed and implemented a 45-minute semi-structured interview guide that was tailored to answer the following key questions:

- What did care for patients receiving LEJR due to fracture look like at the hospital?

- What strategies did the hospital use to try to reduce payments or meet the target price for fractures in the CJR model?

- What non-medical factors did the hospital consider when caring for patients receiving LEJR due to fracture?

- What changes did the hospital make to care for patients receiving LEJR due to fracture during the COVID-19 pandemic?

- What changes did the hospital make in response to the new, fracture specific MS-DRGs (521 and 522) and their inclusion in the CJR model?

Prior to the interview, hospital representatives were asked to complete a brief web-based survey to gather descriptive information about the hospital's standard protocol for hip fracture care, education and discharge planning for hip fracture patients, and any recent efforts to reduce payments for hip fracture patients. We used responses from the pre-interview surveys to tailor the interview protocols.

[1]   Centers for Medicare & Medicaid Services. Comprehensive care for joint replacement model - first annual report appendices. https://innovation.cms.gov/files/reports/cjr-firstannrptapp.pdf. 2018: C3-C10.

[2]   Centers for Medicare & Medicaid Services. Comprehensive care for joint replacement model - second annual report appendices. https://downloads.cms.gov/files/cmmi/cjr-secondannrpt-app.pdf. 2019: E5-E10.

[3]   Centers for Medicare & Medicaid Services. Comprehensive care for joint replacement model - third annual report appendices. https://innovation.cms.gov/data-and-reports/2020/cjr-thirdannrpt-app.pdf. 2020: E5-E11.

[4]   Centers for Medicare & Medicaid Services. Comprehensive care for joint replacement model - fourth annual report appendices. https://innovation.cms.gov/data-and-reports/2021/cjr-py4-ar-app.pdf. 2021: C5-C9.

c.   Interviewee selection criteria

We focused our interview efforts on hospitals that had substantial experience caring for patients with LEJR due to hip fracture. Our sampling universe included CJR mandatory hospitals who had at least five fracture LEJR episodes in 2019 and any fracture LEJR in 2020 to ensure that respondents could offer insight into care for patients with hip fractures under the CJR model during the COVID-19 public health emergency (PHE). From this universe, we randomly sampled 100 hospitals for the telephone interview.

d.   Interviewee recruitment

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.

Our outreach approach was refined based on our prior rounds of telephone interviews. We sent an initial email invitation to each hospital in our sample, 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 and requested that the hospital point of contact complete the pre-interview survey to inform and tailor the interview. On the day-of the interview, we sent a reminder email with a note to review the informed consent attachment and complete the pre-interview survey.

e.   Data collection

We interviewed 39 of the 100 sampled hospitals (39.0%). One interviewer and one note taker conducted the telephone interviews. Notes were taken during telephone interviews, and if the interviewee agreed, the interview was recorded. Recordings were used to verify and enhance interview notes. Notes from telephone interviews were organized and entered into ATLAS.ti software (version 8; Scientific Software Development GmbH, Berlin, Germany) for coding and analysis.

2.   Care coordination survey

Prior qualitative data collection activities associated with the evaluation identified a relationship between participation in the CJR model and increased care coordination activities. To add to the existing knowledge base of CJR model impacts, the care coordination survey aimed to evaluate how care coordination varies across a lower joint replacement episode (pre-surgical, inpatient, and post-discharge or PAC) for CJR participant hospitals.

a.   Survey sample

We used the CJR participant list and episode files based on Medicare Part A institutional claims to identify the universe of mandatory and opt-in hospitals participating in the CJR model in 2019 (n=470). Hospitals with fewer than 20 episodes in 2019 were excluded from the sample (n=75). The final sample included 395 hospitals.

b.   Survey domains

The 25-question survey instrument (Appendix I) was developed in collaboration with CMS. Cognitive testing was performed with six care coordinators prior to fielding and minor refinements were incorporated based on findings. Exhibit C-2 provides information about the survey domains and topics.

Exhibit C-2:   CJR care coordination survey domains and topics of interest

Domain	Topics
Investment in care coordination	▪ Was the care coordinator hired because of the CJR model? ▪ Other additional staff and resources dedicated to care coordination at the hospital
Care coordination activities	▪ How influential was the CJR model on the hospital's decision to implement or enhance different care coordination activities?
Patient selection and transition of care	▪ Frequency with which different patient factors determined discharge destination decision to a SNF or IRF ▪ Frequency with which different issues posed a challenge to patients being safely discharged home ▪ Whether the hospital provided patients with a resource list of community services when they were discharged
Impact of care coordination	▪ How does the hospital measure progress towards their care coordination strategy? ▪ Hospital identified goals for their care coordination activities ▪ Types of outcomes monitored to measure success of the hospital's care coordination strategy
About Respondents	▪ Profession or occupation ▪ Title ▪ Number of years working in care coordination role ▪ Amount of time dedicated to care coordination activities

c.   Data collection

The online survey was fielded from July 28 through August 30, 2020. An updated list of hospital points of contact was obtained from CMS. Contact information for hospital care coordinators was obtained from hospital points of contact, and a link to the survey was emailed to the care coordinators. We used various communication mechanisms to encourage survey response, including an email sent directly from CMS and reminder emails.

The care coordination survey was fielded during the COVID-19 PHE. Thirty-three hospitals declined to participate to focus on distributing their limited care coordination resources to support pandemic needs, and two hospitals declined due to staffing shortages. The final response rate was 50.4% with 199 of the 395 eligible hospitals completing the survey.

II. 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 creation of matched control groups for hospitals in voluntary MSAs, and outlines the additional eligibility criteria for hospitals and episodes.

A. Defining the CJR and control group populations

CMS selected MSAs eligible for CJR participation based on a stratified random sampling methodology in which MSAs were stratified 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 represent what would have happened in CJR-type markets if the model was never implemented (i.e., the counterfactual).

In January 2018, CMS reduced the mandatory participation by about half by allowing all CJR hospitals in the 33 low-payment MSAs and CJR hospitals in the 34 high-payment MSAs that were designated as rural or low-volume a one-time opportunity to remain in the model. The 67 original CJR MSAs were ranked by average historical wage-adjusted episode payment and the top 34 MSAs with the highest payments were required to continue participation in the model (mandatory MSAs), while hospitals in the bottom 33 MSAs were given a one-time opportunity to opt-in (voluntary MSAs). This report covers the first five performance years of the model from April 1, 2016 to September 30, 2021. Our analysis primarily focused on episodes from hospitals that were mandated to participate in PY5 (mandatory analysis). This analysis excluded rural and low-volume hospitals in the mandatory MSAs that were allowed to opt-in to continue participation in CJR. Low-volume hospitals had less than 20 episodes over a three-year historical period (2012 to 2014) and rural hospitals were identified using the FY 2019 IPPS data (section 401 hospitals). In this report, we also present Medicare program savings (MPS) results for the hospitals in the 33 voluntary MSAs that opted to continue participation in PY3 (opt-in hospitals) and those that did not (non-opt-in hospitals).

Exhibit C-3 shows the names and core-based statistical area (CBSA) identification numbers of the CJR and control group MSAs included in the mandatory and voluntary analyses. The MSAs included in the mandatory analysis are starred, while the voluntary MSAs are unstarred. Section II.B provides additional detail about how the control group MSAs were identified and the weights generated for mandatory analyses. Section II.C provides additional detail about how the matched control groups were identified for the voluntary analyses to produce estimates of MPS.

Exhibit C-3: CJR and control group MSAs included in the mandatory and voluntary analyses

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

LEWINGROUP®

CJR		Control	
CBSA ID	MSA name, state	CBSA ID	MSA name, state
33860	Montgomery, AL*	24860	Greenville-Anderson-Mauldin, SC*
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 CMS.

Notes: An asterisk indicates that the MSA was included in the mandatory analysis. MSAs without an asterisk were included in the voluntary opt-in and non-opt-in analyses.

CBSA = core-based statistical area, MSA = metropolitan statistical area.

B. Creation of the analytic weights for the mandatory analysis

1. Average treatment effect of the treated (ATT)

For the original design of the model, the probability of an MSA being selected to participate in the CJR model varied across the strata, with CMS proportionally under-sampling MSAs in the lower average episode payment strata (stratum 1, 2, 5, and 6) and over-sampling MSAs in higher average episode payment strata (stratum 3, 4, 7, and 8). Exhibit C-4 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 C-4:   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 Medicare Program Comprehensive Care for Joint Replacement Payment Model for Acute Care Hospitals Furnishing Lower Extremity Joint Replacement Services; A Final Rule by the Centers for Medicare & Medicaid Services, 80 FR 73273 (November 24, 2015) (codified at 42 CFR 510).

Note: MSA = metropolitan statistical area.

We used an ATT analysis to evaluate the impact of CJR on mandatory hospitals. For this analysis, we constructed the control group using the following steps:

Step 1.   We began with the 104 non-CJR MSAs.

Step 2.   We identified and excluded low-volume and rural hospitals from the 104 non-CJR MSAs because these hospitals were excluded from mandatory participation in the CJR group.

Step 3.   We applied MSA-level weights to the 104 non-CJR MSAs based on the exact probability that the MSA was selected into the 34 mandatory CJR MSAs through the two-step selection process.

To construct the weights in Step 3, we first calculated **the probabilities of the first-stage selection** for each MSA, i.e., 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**, i.e., the probability that the MSA was selected into 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 and so we could not rely on simple proportions.[5] We therefore calculated exact probabilities using combinatorics. We used the exact

---

[5]   They were selected by ranking the original 67 CJR MSAs by historical average episode payment and retaining the top half of the sample (i.e., retaining the 34 MSAs with the highest historical average episode payment).

probabilities to construct MSA-level weights such that the weighted control group was representative of the CJR group. Specifically,

- *Weight for 'mandatory CJR hospitals' = 1*

- *Weight for control group hospitals =*

$$\frac{(probability\ in\ treatment)}{(probability\ in\ control)} = \frac{(probability\ in\ 34\ mandatory\ CJR\ MSAs)}{(probability\ in\ 104\ non\ CJR\ MSAs)}$$

    Note: These are MSA stratum-level weights so all control group hospitals in the same MSA will have the same weight.

We compared the exact probabilities with simulated probabilities that we produced by simulating the two-stage selection process 1,000 times, summing the number of times each MSA was selected into the 34 mandatory CJR MSAs, and dividing the sum by 1,000. The exact probabilities from the combinatorics-based solution and the simulated probabilities are the same (rounded to the 10th of a percent).

Exhibit C-5 shows the analytic weights calculated for control group MSAs included in the mandatory analysis.

Exhibit C-5:   Analytic weights for control group MSAs included in the mandatory analysis

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
3	Ocala, FL	0.67
7	Philadelphia-Camden-Wilmington, PA-NJ-DE-MD	0.56

MSA sampling stratum	MSA	Weight
3	Punta Gorda, FL	0.73
7	Riverside-San Bernardino-Ontario, CA	0.02
7	Salt Lake City, UT	0.05
3	Waterloo-Cedar Falls, IA	0.73
3	Wilmington, NC	0.00
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).

Note: MSA = metropolitan statistical area.

2. *Propensity score weight*

Next, we adjusted the ATT weights to account for CJR participant hospitals shifting a lower share of LEJRs to the hospital outpatient setting. We included outpatient LEJRs in the control group and further adjusted the weights on these outpatient LEJRs to create balance with the CJR group.

TKA was removed from the inpatient only list in January 2018 and THA was removed in January 2020. As a result, Medicare pays for TKAs and THAs performed in the hospital outpatient department; however, the CJR model only includes inpatient LEJRs as episodes. Following the policy changes, both mandatory CJR and control group hospitals began performing TKAs and THAs in the outpatient setting, however mandatory CJR hospitals shifted fewer LEJRs to the outpatient setting. Our analyses indicated that a portion of the CJR inpatient LEJRs would have been outpatient in the absence of the CJR model.[6] As a result of this differential response to the outpatient TKA and THA policies, an appropriate counterfactual for the CJR episodes would need to include patients who would have received their LEJR in the inpatient setting if they had been treated in a CJR hospital, but instead received their LEJR in the outpatient setting because they were treated at a control group hospital.

To construct an appropriate counterfactual, we employed the propensity score weighting (PSW) method and included *all* control outpatient LEJRs in the DiD model, weighted by the hypothetical probability of an outpatient LEJR being inpatient if the hospital had been participating in the CJR model. The probability weights were constructed to ensure that the weighted sum of all control group outpatient LEJRs balances the CJR inpatient LEJRs predicted to have been inpatient LEJRs in the absence of the CJR model. Outpatient LEJRs were not included in the CJR group.

A logit regression was used to model the probability that a LEJR in the CJR or control groups would be performed in the inpatient or outpatient setting. Separate models were created for TKAs

---

6   Centers for Medicare & Medicaid Services. Comprehensive care for joint replacement model - third annual report. https://innovation.cms.gov/data-and-reports/2020/cjr-thirdannrpt.pdf. 2020: 31-37.

and elective THAs. The TKA logit included CJR status, hospital TKA volume in 2017, hospital average length of stay for TKAs in 2017, and all other risk adjustment variables included in our difference-in-differences (DiD) models. The THA logit included the same variables, except hospital elective THA volume in 2019 and hospital average length of stay for elective THAs in 2019. The coefficient on CJR status predicts the proportion of CJR inpatient LEJRs that were inpatient due to the CJR model. A second logit model was run on the CJR-treated inpatient and outpatient LEJRs to predict LEJR setting (inpatient or outpatient). Then, the estimated coefficients from that model were used to predict the probability of a control outpatient LEJR being inpatient had the episode been performed at a CJR hospital. The final weight for each control group LEJR episode was:

$$w * \left(\hat{p} * \frac{N}{\sum \hat{p}} \right)$$

where:

- w is the original sampling weight for the hospital at which the outpatient LEJR was performed.

- $\hat{p}$ is the estimated probability that a control LEJR would have been inpatient had it been performed at a CJR hospital.

- N is the number of control group outpatient LEJRs needed to correct the imbalance in outpatient LEJR shares between CJR and control groups.

- $\sum \hat{p}$ is the sum of all the predicted probabilities for control group LEJRs.

Separate weights were created for TKAs and elective THAs.

C.   Creation of the matched control groups for the voluntary MPS analyses

CJR hospitals located in the 33 voluntary MSAs were given a one-time opportunity in January 2018 to opt to continue participation in the CJR model for PY3 through PY5. We classify these hospitals into two groups: "opt-in CJR hospitals" are hospitals that opted to continue their participation, and "non-opt-in CJR hospitals" are hospitals that did not opt-in and thus their participation ended as of January 1, 2018.

To account for this selection, we constructed a subset group of matched control hospitals to use as a counterfactual when evaluating the impact of the CJR model on each CJR hospital group and estimating MPS. We first took all hospitals located in control MSAs (MSAs eligible but not selected to participate in the CJR model) from sampling strata that also had CJR voluntary MSAs. More specifically, all control MSAs in sampling strata 1, 2, 3, 5, 6, and 7 were included; sampling strata 4 and 8 were used in the mandatory analysis only (Exhibit C-4). Second, we selected specific hospitals located in these control MSAs that resembled the voluntary CJR hospitals on a variety of baseline characteristics. This was performed by separate one-to-one nearest neighbor hospital-level propensity score matching without replacement for opt-in CJR hospitals and non-opt-in CJR hospitals. Each propensity score matching procedure used a logistic regression to estimate propensity scores and included 36 hospital-level covariates calculated using data from our baseline

period.[7] This created a group of matched control hospitals for each of the two groups, specifically 74 opt-in control hospitals and 200 non-opt-in control hospitals, to be used in separate corresponding DiD analyses. Because we created matched control groups for the opt-in and non-opt-in hospitals, we did not need to use analytic weights in these analyses. Additional details pertaining to each matched control group are presented in the following subsections.

1. *Opt-In matched control group balance assessment*

For the 74 opt-in CJR hospitals, we selected 74 matched control group hospitals to serve as a counterfactual in our analyses. Comparing this matched control group to the opt-in CJR group, all but one of the variables used in the propensity score matching procedure had standardized mean differences of less than 0.2.[8,9] The distributions of propensity scores between the matched control group and the opt-in CJR group exhibited common support and appeared similar.[10]

2. *Non-Opt-In matched control group balance assessment*

For the non-opt-in CJR hospitals, we used a caliper in our matching procedure to ensure that the distribution of propensity scores of the non-opt-in CJR hospitals and the matched control hospitals exhibited common support and appeared similar.[11] When matching each non-opt-in CJR hospital with one control hospital, the resulting match had to be within a selected absolute difference (i.e., not exceed a specified threshold) of log-odds propensity score between the two hospitals. The caliper was based on the standard deviation of the estimated log-odds propensity score and assessed among various thresholds to determine the optimal value. We employed a 0.05 caliper, which excluded some non-opt-in CJR hospitals from all our analyses using the matched control group.

---

[7]   The 36 hospital-level covariates included: indicators for hospital ownership, number of hospital beds, total TKA episode volume, total THA episode volume, total LEJR MS-DRG 469 episode volume, total LEJR MS-DRG 470 episode volume, indicator for participation in BPCI LEJR, percent of total LEJR volume that was in BPCI, indicators for Census Division, average HCC score, average age, percent of LEJR patients in age categories (20-64, 65-79, 80+), percent of LEJR patients that were female, percent of LEJR patients in race and ethnicity categories, percent of LEJR patients eligible for Medicaid, percent of LEJR patients with disability excluding ESRD, percent of LEJR patients flagged with obesity, percent of LEJR patients flagged with hypertension, percent of LEJR patients flagged with dementia, and percent of LEJR patients with prior care use six months prior to anchor hospitalization (ACH stay, IRF stay, SNF stay, HH use, any prior care).

[8]   The indicator for Census South Atlantic Division had a standardized mean difference of 0.24 between the opt-in CJR and the matched control group. This was driven by there being 9 opt-in CJR hospitals and only 4 matched control hospitals. Given the similarities in all other matching variables, we do not think this slight geographical imbalance is of concern.

[9]   Stuart, E.A. (2010). Matching methods for causal inference: A review and a look forward. Statistical science: a review journal of the Institute of Mathematical Statistics, 25(1), 1.

[10] The distributions of the log odds of the propensity score between the opt-in CJR hospitals and the matched control hospitals resulted in failing to reject the null hypothesis of the Kolmogorov–Smirnov test that the distributions were equal (p=0.65).

[11] When using the caliper, the distributions of the log odds of the propensity score between the non-opt-in CJR hospitals and the matched control hospitals resulted in failing to reject the null of the Kolmogorov–Smirnov test that the distributions were equal (p=0.46).

For the remaining 200 non-opt-in CJR hospitals, we selected 200 matched control group hospitals to serve as a counterfactual in our analyses.[12] With the caliper, all matching variables had standardized mean differences within +/- 0.2.

3. Overlap of voluntary and mandatory control groups

The propensity score matching procedures were performed separately and independently for each of the two groups of CJR hospitals in voluntary MSAs (opt-in and non-opt-in). As a result, the matching procedures considered the same set of potential control group hospitals and were permitted to choose the same individual hospitals. This methodological choice was made based on conceptual factors and assessment of empirical evidence of the quality of the matched control groups. Of the hospitals chosen in the two matched control groups, 31 control hospitals were included in both groups.

Moreover, since the analytic weights used for the analysis of mandatory CJR hospitals included control MSAs from strata 3 and 7, the matching procedures also considered some control hospitals that were included in the mandatory control group. This methodological choice was made to account for these MSA strata not having a certain chance of being hypothetically selected as a "mandatory" MSA. Thus, 18 control hospitals chosen in the matched control group for opt-in CJR hospitals and 32 control hospitals chosen in the matched control group for the non-opt-in CJR hospitals are included in the mandatory control group with nonzero analytic weights.

D. Additional eligibility criteria for hospitals and episodes

1. Hospital criteria

For inclusion in the analysis, hospitals had to be acute care hospitals (ACH) paid under the IPPS that performed LEJR for Medicare beneficiaries in the baseline or intervention periods.

2. Episode definition

For both the CJR and control group populations, the beginning of an episode is 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 the two new MS-DRGs for LEJR due to hip fracture (521 with MCC and 522 without MCC) as episode triggers.[13] The end of the episode is 90 days after the anchor hospital discharge.

---

[12] Two non-opt-in CJR hospitals did not have LEJR episode volume in the baseline and thus were excluded from our analyses.

[13] New MS-DRGs were adopted by CMS for LEJRs due to hip fracture in October 2020 (521 for LEJR due to hip fracture with MCC; 522 for LEJR due to hip fracture without MCC). Previously, hip fracture episodes were discharged under MS-DRGs 469 or 470 and were identified as having a hip fracture based on ICD diagnosis codes.

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 (i.e., not enrolled in any managed care plan or covered under other health plans); and
- not eligible for Medicare based on end-stage renal disease (ESRD).

As specified in the Final Rule, episodes were cancelled in the CJR model and excluded from the analysis if:

- the patient no longer met the eligibility criteria described in the preceding paragraph;
- 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, a Medicare Shared Savings Program ACO track 3, or a Comprehensive ESRD Care Model ACO with downside risk;[14] or
- the episodes were attributed to the Bundled Payments for Care Improvement initiative.[15]

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 (age, sex, and six months of Medicare FFS enrollment history prior to the LEJR hospital admission).

We also created outpatient LEJR episodes for inclusion in the control group, as described above in Section II.B.2. Beginning in January 2018, CMS removed TKA from the inpatient only list, allowing Medicare coverage for TKAs provided in the hospital outpatient setting. THA was removed in January 2020. Evidence suggests that the CJR model influences the choice of inpatient or outpatient setting, which would bias impact estimates that are based only on inpatient LEJR episodes that are included under the CJR model. (Annual Report 3 includes additional information

---

[14] 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

[15] Episodes initiated at CJR participant hospitals could be attributed to a physician group practice (PGP) participating in the Bundled Payments for Care Improvement initiative or to skilled nursing facilities, inpatient rehabilitation facilities, long-term care hospitals or home health agencies participating in the Bundled Payments for Care Improvement Initiative Model 3.

about outpatient TKA and the CJR model.[16]) Therefore, we also include outpatient LEJR episodes in the control group and apply a weight based on their probability of being an inpatient LEJR in the absence of CJR to obtain impact estimates of the CJR model. For the outpatient LEJRs, the beginning of the episode was triggered by an LEJR performed in the outpatient department of a CJR participating or control group hospital (CPT code 27447 [for TKA] or 27130 [for THA] assigned to C-APC 5115 with status indicator "J1" in Part B institutional claims). The end of the episode is 90 days after the outpatient procedure and beneficiaries had to meet and maintain the CJR eligibility criteria throughout the episode to be included in the analysis.

---

[16]   Centers for Medicare & Medicaid Services. Comprehensive care for joint replacement model - third annual report. https://innovation.cms.gov/data-and-reports/2020/cjr-thirdannrpt. 2020: 31-37.

III. Impact of the CJR Model on Claims and Assessment-based Outcomes

A. Measures of impact on payments, utilization, and quality

In this section we present the episode-level outcome measures that were constructed using Medicare FFS claims to assess the impact of the CJR model on Medicare payments, utilization, and quality (Exhibit C-6).

Exhibit C-6:   Claims-based payment, utilization, and quality measures

Measure category	Measure name/description
Medicare payments[a]	Total Medicare standardized allowed amounts included in the episode, inpatient anchor hospitalization through the 90-day PDP
	Medicare standardized allowed amounts included in the inpatient anchor hospitalization
	Medicare standardized allowed amounts per episode, by service, 90-day PDP[b]
	Medicare standardized allowed amounts, 30-Day PEP[c]
Utilization	First post-acute discharge was to IRF
	First post-acute discharge was to SNF
	First post-acute discharge was to HHA
	First post-acute discharge was home without HHA
	Any HHA visits, 90-day PDP
	Number of IRF days, 90-day PDP[d]
	Number of SNF days, 90-day PDP[d]
	Number of HHA visits, 90-day PDP[d]
	Number of outpatient PT/OT visits[d]
Quality	Unplanned readmissions, 90-day PDP
	Emergency department visits, 90-day PDP
	All-cause mortality, inpatient stay and 90-day PDP[e]
	Incidence of any complications, 90-day PDP[f]

Source: All measures are constructed from Medicare fee-for-service claims data.

Notes: HHA = home health agency, IRF = inpatient rehabilitation facility, OT = occupational therapy, PAC = post-acute care, PDP = post-discharge period, PEP = post-episode period, PT = physical therapy, SNF = skilled nursing facility.

[a]   Payments are the standardized Medicare allowed amounts. Standardization removes wage adjustments and other Medicare payment adjustments. Allowed amounts include beneficiary cost sharing.

[b]   Services include inpatient readmissions, IRF, SNF, HHA (Parts A and B), and services covered under Medicare Part B.

[c]   Services include all health care services covered under Medicare Part A and Part B.

[d]   The eligible sample for PAC days and visits is among those with any use.

[e]   Under the CJR model, death during the anchor hospitalization or 90-day PDP cancels the episode. Therefore, to estimate the all-cause mortality rate, this analysis includes CJR and control group episodes as well as beneficiary admissions at CJR and control group hospitals that would have been identified as episodes if the beneficiaries had not died during the anchor hospitalization or 90-day PDP.

[f]   THA/TKA complications is measured among elective episodes only.

B. Measures of unintended consequences

Our evaluation of unintended consequences of the CJR model focused on changes in patient mix. Exhibit C-7 lists the patient characteristics from claims and enrollment data that we monitored.

While the impact analysis on payment, utilization, and quality controlled for changes in these patient characteristics, we also monitored changes in these characteristics separately to directly examine changes in patient mix.

Exhibit C-7:   Measures of patient mix

Type of unintended consequence	Measure name/description
Changes in patient mix	Age
	Sex
	Race and ethnicity
	Medicaid eligibility
	Disability, no ESRD
	Congestive heart failure
	HCC score
	Dementia
	Obesity
	Hypertension
	Diabetes
	Prior utilization (in the six months prior to the anchor hospitalization) ▪ Inpatient ACH stay ▪ IRF stay ▪ SNF stay ▪ Home health use ▪ Any prior care[a]

Source: Patient mix measures are constructed from Medicare fee-for-service claims and beneficiary enrollment data.

Notes: ACH = acute care hosptial, ESRD = end-stage renal disease, HCC = hierarchical condition category, IRF = inpatient rehabilitation facility, SNF = skilled nursing facility.

   [a] Any prior care includes inpatient hospital, psychiatric hospital, emergency department, skilled nursing facility, inpatient rehabilitation facility, home health, long-term care hospital, and hospice during the six months prior to anchor hospitalization.

C. Analytic methodology

While the CJR and control group populations are overall quite similar in terms of market, hospital, and patient characteristics,[17] there may be unobserved differences that impact outcomes. To control for both observed and unobserved differences and to isolate the impact of the CJR model on outcomes, we used a DiD regression approach supplemented by risk adjustment.

1. DiD estimator

The DiD approach quantifies the impact of the CJR model by comparing changes in outcomes between the baseline and intervention periods 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.[18]

---

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

[18] While the DiD model controls for unobserved heterogeneity that is fixed over time, it does not control for unobserved heterogeneity that varies over time.

a. Baseline period

The baseline period for our evaluation encompasses 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 follows the model timeline stipulated in the CJR Final Rules. The first performance year included episodes starting on or after April 1, 2016. PY5 was originally set to end on December 31, 2020.[19] However, due to the COVID-19 public health emergency, CMS extended PY5 by nine months.[20] The intervention period for this report includes episodes ending through September 30, 2021.

The DiD model uses an outcome measure, *Y*, and estimates the differential change in *Y* for beneficiaries receiving care from CJR participant hospitals between the baseline and the intervention periods relative to that same change for beneficiaries receiving care from hospitals in the control group.

To illustrate the DiD approach, we define:

- $Y_{i,k,t}$ is the outcome for the ith episode with an LEJR at hospital k in period t ($t = 1$ during the CJR intervention quarters and zero otherwise)

- $CJR_{i,k}$ is an indicator that takes the value of 1 if the ith episode was initiated by a CJR participant hospital k and takes the value of 0 otherwise

- $X_{i,k,t}$ are hospital, geographic, and patient characteristics in period t

- E[Y|t, CJR, X] is the expected value of outcome measure Y conditional on values of t, *CJR*, and *X*

The DiD estimator is:

$$DiD = [E(Y \mid t=1, CJR = 1, X) - (E(Y \mid t=0, CJR = 1, X)] - [E(Y \mid t=1, CJR = 0, X) - (E(Y \mid t=0, CJR = 0, X)] \qquad \textbf{(1)}$$

To illustrate the calculation of the DiD, consider the linear model listed below:

$$Y_{i,k,t} = b_0 + b_1 \cdot t + b_2 \cdot CJR_{i,k} + b_3 \cdot CJR_{i,k} \cdot t + X_{i,k,t}' \cdot B + u_{i,k,t} \qquad \textbf{(2)}$$

- The value of coefficient b_1 captures 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 captures the relative differences in outcomes between CJR and control group episodes.

---

[19] CMS. Medicare Program; Comprehensive Care for Joint Replacement Payment Model for Acute Care Hospitals Furnishing Lower Extremity Joint Replacement Services: final rule (42 CFR Part 510). *Fed Regist*. 2015; 80(226): 73273-73554.

[20] CMS. Additional Policy and Regulatory Revisions in Response to the COVID–19 Public Health Emergency (85 FR 71142). *Fed Regist*. 2020; 85(216): 71142-71205.

- Coefficient b_3 determines 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 represents the DiD estimator.

- The vector of coefficients B measures the differential effects of risk factors (X) on the outcome variable.

To calculate separate DiDs for each of the five performance years during the intervention period, Equation 2 was modified to include five time period indicators t_1 (equals 1 during PY1 intervention period and zero otherwise), t_2 (equals 1 during PY2 intervention period and zero otherwise), t_3 (equals 1 during PY3 intervention period and zero otherwise), t_4 (equals 1 during PY4 intervention period and zero otherwise), and t_5 (equals 1 during PY5 intervention period and zero otherwise).

$$Y_{i,k,t} = b_0 + b_1 \cdot t_1 + b_2 \cdot t_2 + b_3 \cdot t_3 + b_4 \cdot t_4 + b_5 \cdot t_5 + b_6 \cdot CJR_{i,k} + b_7 \cdot CJR_{i,k} \cdot t_1$$
$$+ b_8 \cdot CJR_{i,k} \cdot t_2 + b_9 \cdot CJR_{i,k} \cdot t_3 + b_{10} \cdot CJR_{i,k} \cdot t_4 + b_{11} \cdot CJR_{i,k} \cdot t_5 + X_{i,k,t}' \cdot B + u_{i,k,t} \quad \textbf{(3)}$$

- Coefficient b_7 determines the differential in outcome Y experienced by beneficiaries receiving services from CJR providers during the CJR **PY1** intervention period relative to control group episodes in the **PY1** intervention period and represents the DiD estimator for **PY1**.

- Coefficient b_8, b_9, b_{10}, and b_{11} represent the DiD estimators for **PY2, PY3, PY4, and PY5 respectively**.

Finally, to calculate the DiD estimate for outcome measures that were risk-adjusted with non-linear 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. In these cases, the standard errors were computed using the Delta method.[21] For all DiD models, statistical significance was assessed at the 10% level.

This approach was used for mandatory and voluntary analyses. For the mandatory analysis, we applied the analytic weight described above in Section II.B. We used the matched control group, described in Section II.C, for the voluntary analysis, which did not require an analytic weight.

c. Assumptions of DiD estimators

One critical assumption of an unbiased DiD estimate is that the treatment and control group outcomes follow parallel trends for the outcome of interest during the baseline period. Another assumption is that these parallel trends would have remained the same in the period when the policy is actually implemented in the absence of the policy intervention. While the first assumption

---

[21] The delta method expands a function of a random variable about its mean, usually with a Taylor approximation, and then takes the variance. Specifically, if $Y = f(x)$ is any function of a random variable X, we need only calculate the variance of X and the first derivative of the function to approximate the variance of Y. Let μ_x be the mean of X and $\dot{f}(x)$ be the first derivative, a Taylor expansion of $Y = f(x)$ about μ_x gives the approximation: $Y = f(x) \approx f(\mu_x) + \dot{f}(\mu_x)(x - \mu_x)$. Taking the variance of both sides yields: $Var(Y) = Var(f(X)) \approx [\dot{f}(\mu_x)]^2 Var(X)$. For example, suppose $Y = X^2$. Then $f(x) = X^2$ and $\dot{f}(x) = 2x$, so that $Var(Y) \approx (2\mu_x)^2 Var(X)$.

can be tested if sufficient baseline data on the CJR and control groups are available, the second assumption is untestable.

We evaluated the parallel trends assumption two ways: visually inspected trends for all outcomes; and statistically tested that the CJR and control group outcomes follow parallel trends during the baseline period. 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 is evidence to reject the parallel trend assumption. We considered outcomes to fail parallel trends if we reject the null hypothesis of seemingly parallel trends at the 10% significance level for both tests.

For the joint F-test, we report the p-value of an F-test that tests if the differential between the CJR and control groups is jointly equal across discrete four-quarter time periods. We included dummy variables for each of the three baseline years; interaction terms between the CJR group indicator and each of the year dummies, along with all the risk-adjustment variables that we include in the DiD models (described in the Section III.C.2).

The joint F-test model is:

$$Y_{i,k,t} = b_0 + b_1 \cdot Year_{1,i} + b_2 \cdot Year_{2,i} + b_3 \cdot Year_{3,i} + b_4 \cdot Year_{1,i} \cdot CJR_k + b_5 \cdot Year_{2,i} \cdot CJR_k + b_6 \cdot Year_{3,i} \cdot CJR_k + X_{i,k}' \cdot B + u_{i,k,t}$$

where:

- $Y_{i,k,t}$ is the outcome for the ith episode with an LEJR at hospital k in the baseline period in year t.

- $Year_{i,t}$ is an indicator that takes the value of 1 if the ith episode was initiated during year t of the baseline period and takes the value of 0 otherwise

- $CJR_{i,k}$ is an indicator that takes the value of 1 if the ith episode was initiated by a CJR participant hospital k and takes the value of 0 otherwise

- $X_{i,k}$ are hospital, geographic, and patient characteristics in the baseline period

and the test is:

$$H_0 : b_4 = b_5 = b_6$$

$$H_1 : b_4 \neq b_5, \text{ or } b_4 \neq b_6, \text{ or } b_5 \neq b_6$$

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 include in the DiD models.

The linear test model is:

$$Y_{i,k,t} = b_0 + b_1 \cdot Quarter_{i,t} + b_2 \cdot CJR_k + b_3 \cdot Quarter_{i,t} \cdot CJR_k + X_{i,k,t}' \cdot B + u_{i,k,t}$$

where:

- $Y_{i,k,t}$ is the outcome for the ith episode with an LEJR at hospital k in the baseline period in quarter t.

- $Quarter_{i,t}$ is an indicator that takes the value of 1 if the ith episode was initiated during quarter t of the baseline period and takes the value of 0 otherwise

- $CJR_{i,k}$ is an indicator that takes the value of 1 if the ith episode was initiated by a CJR participant hospital k and takes the value of 0 otherwise

- $X_{i,k}$ are hospital, geographic, and patient characteristics in the baseline period

and the test is:

$$H_0: b_3 = 0$$

$$H_1: b_3 \neq 0$$

For mandatory CJR hospitals, the following outcomes failed both the linear and joint F-tests of parallel trends:

- HHA payments ($p<0.05$ for the joint test and $p<0.01$ for the linear test for all LEJR and elective LEJR)

- 30-day post-episode payments ($p<0.01$ for the joint test and $p<0.05$ for the linear test for all LEJR and elective LEJR)

- First PAC SNF ($p<0.10$ for both joint and linear tests for fracture LEJR)

- Number of outpatient PT/OT visits ($p<0.10$ for the joint test and $p<0.05$ for the linear test for all LEJR)

Results for the parallel trends tests are included in Appendix K.

2. 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. All outcomes were risk adjusted for the episode's hip fracture status, procedure type (hip or knee), and presence or absence of a major complication or comorbidity, defined by MS-DRG. To control for participation in other Medicare initiatives, we used a dummy variable that indicated whether the beneficiary was in the Medicare Shared Savings Program (MSSP), Pioneer ACO

Model, or Next Generation ACO Model during the episode.[22] To control for prior health

conditions, we used HCC indicators for the 12 months preceding the anchor hospitalization,[23] 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 prior care use in the following settings: acute care IPPS hospital, emergency department, LTCH, SNF, IRF, hospice, other Part A inpatient, HHA, and custodial nursing facility. We used the MDS assessment data to create a measure of prior custodial nursing facility use in the six months prior to the episode (see call out box).

> **Measuring prior nursing facility use**
>
> During the COVID-19 PHE, MDS assessment data reporting and transmission requirements were waived by CMS (later reinstated in May 2021). Consequently, data on whether a beneficiary had a custodial nursing facility stay or SNF stay in the six months prior to the episode were missing during this period. We imputed values of prior nursing facility use during the reporting pause using beneficiary characteristics, including prior health care use, anchor MS-DRG, age, sex, disability (not ESRD), Medicaid eligibility, chronic conditions (indicators of diabetes, obesity, tobacco use, and HCC flags), and race and ethnicity. The misclassification rate was 0.5%. We retained observed nursing facility stays during the reporting pause and pre-PHE data remained unchanged.

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).[24] To address regional effects of the COVID-19 pandemic, we included cumulative county-level case rates, as well as county-level seven-day moving average case rate at the start of the episode and 21 days after the start of the episode.

We also controlled for provider characteristics that might be related to the outcomes of interest, such as hospital bed count, for-profit status, and previous Bundled Payments for Care Improvement initiative LEJR experience and previous Bundled Payments for Care Improvement initiative experience in a clinical episode other than LEJR. In October 2018, the Bundled Payments for Care Improvement Advanced initiative began. This CMMI model also includes LEJR as a clinical episode and aims to reduce payments, while maintaining or improving quality. CJR participant hospitals could not participate the in Bundled Payments for Care Improvement

---

[22]   Beneficiaries with episodes during or after July 2017 that were aligned with MSSP track 3, Next Generation ACO, or Comprehensive End Stage Renal Disease Care Model and were excluded from the CJR model.

[23]   The Hierarchical Condition Category (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. The HCC indicators in the risk adjustment model included: sepsis, different types of cancer, diabetes, obesity, malnutrition, rheumatoid arthritis, coagulation defects, dementia, drug/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.

[24]   In PY5, 1.1% of episodes in the mandatory CJR group had a COVID-19 diagnosis in the 30-days prior to the anchor hospitalization, compared to 1.0% in the control group. During the anchor hospitalization, 3.3% of episodes in the mandatory CJR group had a COVID-19 diagnosis, compared to 2.9% in the control group.

Advanced initiative for LEJR clinical episodes; however, hospitals and surgeons in the control group could participate. We found that 46% of mandatory control group episodes that started on or after October 1, 2018 were attributed to the Bundled Payments for Care Improvement Advanced initiative. To account for contamination in our control group by this other CMMI model, we included indicator variables that identify control group LEJR episodes performed by surgeons or at hospitals participating in the Bundled Payments for Care Improvement Advanced model. One dummy variable accounts for inpatient LEJR attributed to Bundled Payments for Care Improvement Advanced in model years 1 and 2 (October 2018 through December 2019). In January 2020, outpatient TKA became an episode in the Bundled Payments for Care Improvement Advanced initiative. The second dummy variables accounts for inpatient LEJR and outpatient TKA attributed to the model in year 3 and later (on or after January 2020).

While the same demographic and enrollment status indicators were included for all outcomes, we considered alternative aggregation levels to control for prior care use, prior health conditions, and regional characteristics (Exhibit C-8). To assess different specifications, we split the sample into a model development and a validation sample and estimated each model using data from the model development sample. We then evaluated the models' goodness of fit (Akaike Information Criterion (AIC), Bayesian Information criterion (BIC) criteria, and R-square) in the model development sample and their predictive performance in the validation sample.

Exhibit C-8:   Predictive risk factors used to risk-adjust claims-based outcomes

Domain	Variables
Characteristics of the procedure	▪ Anchor MS-DRG ▪ 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 ESRD) ▪ Attribution to Medicare Shared Savings Program, Pioneer ACO Model, or Next Generation ACO Models during the CJR episode
Prior health conditions	▪ CMS-HCC version 21 indicators from qualifying services and diagnoses (those meeting a threshold of at least 1%) from claims and data for 12 months preceding the anchor hospitalization ▪ Obesity indicator ▪ Diabetes indicator ▪ Hypertension indicator ▪ Tobacco use indicator

Prior use	▪ Prior care use (any acute care inpatient, ED visits, IRF, SNF, HHA, hospice, other Part A inpatient, LTCH, and custodial nursing facility service) variables used in risk adjustment varied by model[a]
	  • Binary indicators for any care use in the six months preceding the start of the episode
	  • Binary indicators for any care use in the one month preceding the start of the episode
	  • Number of days of care use in the six months preceding the start of the episode

Domain	Variables
Geography	▪ State indicators
COVID-19	▪ County-level cumulative COVID-19 infection rate at the start of the episode ▪ County-level seven-day moving average of the new COVID-19 infection rate at the start of the episode ▪ County-level seven-day moving average of the new COVID-19 infection rate 21 days after the start of the episode ▪ 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 ▪ Bundled Payments for Care Improvement LEJR experience ▪ Bundled Payments for Care Improvement experience in a clinical episode other than LEJR ▪ LEJR performed by surgeons or at hospitals participating in the Bundled Payments for Care Improvement Advanced 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 POS, FY 2016 CMS Annual IPPS, USAFacts Coronavirus Live Map, CMS Master Data Management, Bundled Payments for Care Improvement initiative participant list, and Bundled Payments for Care Improvement Advanced initiative participant list.

Notes: ACO = accountable care organization, COVID-19 = coronavirus disease 2019, ED = emergency department, ESRD = end-stage renal disease, FY = fiscal year, HCC = hierarchical condition category, HHA = home health agency, IPPS = inpatient prosepctive payment system, IRF = inpatient rehabilitation facility, LEJR = lower extremity joint replacement, LTCH = long-term care hospital, MS-DRG = Medicare Severity-Diagnosis Related Group, POS = provider of services, SNF = skilled nursing facility.

[a]   The optimal specification for each prior use variable was chosen using the goodness of fit criteria 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/OT visits.

3.   Model types

We used a variety of models including logistic, Poisson, multinomial logit, OLS regressions, and two-part models (Exhibit C-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 (i.e., all claims-based quality of care measures). A multinomial logit model was applied to first-discharge setting. OLS models were estimated for the continuous total number of days or visits measures (e.g., number of SNF days, number of IRF days, number of HHA visits, and number of outpatient PT/OT visits), as well as total episode payments and Part B payments. Two-part models were favored for payment outcomes where more than 5% of individuals had zero payments for the particular outcome. These payment outcomes included the individual Part A payments that exhibited zero-mass and skewness.

Exhibit C-9:　Outcomes by model type

Model type	Outcomes
Ordinary least squares (OLS)	▪ Total episode payments ▪ Part B payments ▪ Number of IRF days ▪ Number of SNF days ▪ Number of HHA visits ▪ Number of outpatient PT/OT Visits
Two part models (Probit/OLS)	▪ Readmission payments ▪ IRF payments ▪ SNF payments ▪ HHA payments ▪ 30-day PEP 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 home health
Logistic	▪ Any HHA visits ▪ Unplanned readmission ▪ Emergency department visit ▪ Complications, among elective episodes ▪ All-cause mortality

Note:　　HHA = home health agency, IRF = inpatient rehabilitation facility, OLS = ordinary least squares, OT = occupational therapy, PEP = post-episode period, PT = physical therapy, SNF = skilled nursing facility.

Estimates from the multivariate regression models were used to construct model-predicted outcomes under two scenarios (baseline and 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 in this report is all CJR beneficiaries during the baseline and intervention period. Given the design of the CJR model (randomly sampling MSAs to participate), we accounted for clustering at the MSA level in the estimation of our standard errors in all regression models for mandatory hospitals. In our regression models for opt-in and non-opt-in hospitals, we accounted for clustering at the hospital level in the estimation of our standard errors because the decision whether or not to continue participation in PY3 was at the hospital level.

D. Outpatient LEJR descriptive methods

Medicare now covers LEJRs performed in the hospital outpatient setting.[25,26] CMS removed TKA from the inpatient only list, effective January 2018, and THA was removed starting January 2020. We monitored the occurrence of outpatient LEJR in CJR and control group hospitals and MSAs from January 1, 2018 through July 3, 2021.[27]

We identified LEJRs performed on outpatients using Part B claims data (CPT code 27447 for TKA or 27130 for THA assigned to C-APC 5115 with status indicator "J1") and LEJRs performed on inpatients using Part A claims data (MS-DRG 469 or 470 with ICD codes on the claim indicating a knee or hip procedure). We excluded LEJR discharges that did not meet CJR episode eligibility and THAs performed due to hip fracture.

We calculated percent of LEJRs performed as outpatients overall and by quarter for the mandatory CJR and control groups for TKAs and elective THAs separately. We divided the number of outpatient LEJRs meeting episode eligibility by the sum of all LEJRs meeting episode eligibility (LEJRs performed on inpatients and outpatients).

---

[25] CMS. Medicare Program: Hospital Outpatient Prospective Payment and Ambulatory Surgical Center Payment Systems and Quality Reporting Program. https://s3.amazonaws.com/public-inspection.federalregister.gov/2017-23932.pdf. 2017.

[26] CMS. Medicare Program: Changes to Hospital Outpatient Prospective Payment and Ambulatory Surgical Center Payment Systems and Quality Reporting Programs; Revisions of Organ Procurement Organizations Conditions of Coverage; Prior Authorization Process and Requirements for Certain Covered Outpatient Department Services; Potential Changes to the Laboratory Date of Service Policy; Changes to Grandfathered Children's Hospitals Within-Hospitals; Notice of Closure of Two Teaching Hospitals and Opportunity To Apply for Available Slots. https://www.govinfo.gov/content/pkg/FR-2019-11-12/pdf/2019-24138.pdf. 2019.

[27] To match the episode inclusion criteria for the DiD analyses, inpatient and outpatient LEJR discharges were included if the 90-day post-discharge period occurred on or before September 30, 2021. A patient discharged on July 3 would have a 90-day post-discharge period ending on September 30.

IV. Savings to Medicare due to the CJR model

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

Medicare savings = Change in non-standardized paid amounts – Reconciliation payments

A. Change in non-standardized paid amounts

The change in non-standardized paid amounts was based on estimates from a DiD model of per-episode standardized paid amounts. The DiD estimates were multiplied by negative one and converted to non-standardized paid amounts using a ratio of non-standardized to standardized Medicare paid amounts from CJR intervention episodes (Exhibit C-10). This method produced a per-episode estimate of the change in non-standardized paid amounts. The total change in non-standardized paid amounts was produced by multiplying the per-episode estimate by the total number of episodes.

Exhibit C-10: Ratios of non-standardized to standardized Medicare paid amounts by hospital group

Time period	Mandatory hospitals	Opt-in hospitals	Non-opt-in hospitals
Baseline	1.032	1.035	1.000
PY1	1.032	1.037	1.002
PY2	1.032	1.040	1.003
PY3	1.036	1.037	
PY4	1.042	1.045	
PY5.1	1.056	1.051	
PY5.2	1.073	1.071	
Cumulative	1.043	1.044	1.002

Source: CJR evaluation team analysis of Medicare claims and enrollment data for episodes initiated in 2012 through 2014 that ended between April 2012 and March 2015 (baseline) and episodes initiated during or after April 2016 that ended by September 2021 (intervention).

Notes: The ratio is calculated as the average non-standardized (actual) paid amounts divided by the average standardized paid amounts for episodes. The anchor payment (MS-DRG payment for inpatient episodes) was subtracted from the total episode payment before calculating the ratio.

MS-DRG = Medicare Severity-Diagnosis Related Group, PY = performance year.

B. Reconciliation payments

Reconciliation payments are defined as total payments made to CJR participants by Medicare net of repayments from CJR participants to Medicare. Reconciliation payments can be positive or negative. In the program literature, they are often referred to by the term "net payment reconciliation amounts" or "NPRA." These data were provided by the CMS CJR payment contractor. Reconciliation payments per episode were calculated by dividing total reconciliation payments by the total number of CJR episodes.

C. Hospital group estimates versus cumulative estimates

We reported estimates for three different hospital groups: mandatory CJR hospitals (excluding those with a low-volume or rural exemption), opt-in hospitals in voluntary MSAs, and non-opt-in hospitals in voluntary MSAs.[28]

For each hospital group, a comparison group of episodes from control hospitals was constructed, and a DiD model was used to produce an estimate of per-episode reductions in standardized paid amounts.[29]

We also reported estimates for all hospital groups combined. We could have added together the hospital group estimates or constructed a weighted average of the per-episode estimates from each hospital group. However, these approaches would have led to overly conservative (too wide) confidence intervals and ranges. Instead, we pooled the three regressions together into a single overarching model. This allowed us to construct an accurate confidence interval for the weighted average of per-episode reductions in standardized payments.

D. Performance year estimates versus total estimates

We reported Medicare savings estimates for all five performance years combined, and on a per-performance year basis. The performance year estimate was derived from a DiD model that compared episode payments in a given performance year to episode payments during the baseline period. Thus, for instance, the PY2 estimate was determined by comparing the change in standardized payments per episode between PY2 and the baseline period in CJR hospitals to that same change in control hospitals.

Estimates of the total savings to Medicare over multiple performance years for a specific hospital group were constructed from the performance year estimates. We constructed a weighted average of the performance year estimates, with the weights reflecting the proportion of all episodes that occurred in a given performance year. We also estimated a confidence interval for the weighted average, allowing for the construction of our estimated ranges on total savings.

Estimates of the total savings to Medicare including all performance years and all hospitals were constructed by pooling hospital groups as described above, and then constructing a weighted average of the performance year estimates from the pooled regression.

E. Sensitivity Analyses

We also used episode-level NPRA data to estimate Medicare savings under two hypothetical alternative policy responses to the COVID-19 PHE – one in which there was no policy response to

---

[28] Twenty-two opt-in hospitals chose to discontinue their participation in the CJR model in PY5.2. Their participation ended December 2020. Episodes from these hospitals were excluded from the PY5.2 DiD.

[29] To construct an appropriate counterfactual, we employed the propensity score weighting (PSW) method and included *all* control outpatient LEJRs in the DiD model, weighted by the hypothetical probability of an outpatient LEJR being inpatient if the hospital had been participating in the CJR model.

the COVID-19 PHE, and one in which hospitals were shielded from downside risk at the *hospital* level, rather than at the episode level as under the actual COVID-19 PHE policy to remove downside risk.

For the no response scenario, we calculated NPRA for PY5 episodes under the standard reconciliation guidelines (i.e., guidelines used prior to the COVID-19 PHE). We calculated the difference between actual episode spending and target spending for each episode category (MS-DRG and hip fracture status). Then we summed the differences to obtain hospital-level NPRA. Hospitals with positive NPRA received reconciliation payments because, on average, their total payments were less than their quality-adjusted target price. Hospitals with negative NPRA had to repay to Medicare the average difference between the actual episode payments and quality-adjusted target prices (i.e., repayment).

For the hospital-level downside risk waiver scenario, we calculated NPRA for PY5 as described above in the no response scenario separately for episodes initiated before January 30th, 2020, and for episodes initiated on or after January 31st, 2020. Then, for episodes initiated on or after January 31st, 2020, we set final NPRA to zero for any hospital with repayment responsibility.

After obtaining hospital-level NPRA under each scenario, we calculated total NPRA across all hospitals and used the same estimates of payment reductions that entered the main calculation of Medicare savings. It should be noted that this does not account for the possibility that hospitals would have changed their behavior in response to the hypothetical alternative policy responses to the COVID-19 PHE.

F. Considerations

We do not include the low-volume and rural hospitals located in the 34 mandatory MSAs in the analysis of opt-in hospitals in voluntary MSAs because low-volume and rural hospitals differ in important ways that are likely to affect performance in the model. Further, unlike the voluntary hospitals, the low-volume and rural hospitals are located in MSAs with higher average historical payments and the majority of hospitals in these mandatory MSAs are participating in the CJR model. Since an analysis of low-volume and rural hospitals would also need to account for their ability to select to continue in the model, we do not include them in the analysis of the CJR hospitals in the 34 mandatory MSAs that were continuously required to participate throughout the entire model. Producing a separate impact estimate for this subgroup would be a challenge because it would require constructing an appropriate comparison group. This group of hospitals is small (39 low volume hospitals and 37 rural hospitals, with 20% opting to continue participation in PY3), and hospitals were incentivized by the CJR model to reclassify to rural by offering rural hospitals lower stop-loss limits (e.g., 5% in PY4 compared to 20% for all other hospitals). The CJR model may have influenced hospitals decision to reclassify as rural and this same incentive to reclassify to rural was not present in the control group.

Rather than producing separate impact estimates for this group, we performed sensitivity analyses in the fourth annual report, which suggested that inclusion of the low volume and rural hospitals

located in mandatory MSAs would not have changed our Medicare program savings results and conclusions.[30] Thus, we excluded low-volume and rural hospitals in mandatory MSAs from our main estimates of Medicare savings in this report as well.

---

[30] Centers for Medicare & Medicaid Services. Comprehensive care for joint replacement model - fourth annual report appendices. https://innovation.cms.gov/data-and-reports/2021/cjr-py4-ar-app.pdf. 2021: E1-E3.

V. Patient Selection

A. Analyses of a composite measure of patient characteristics

As multiple patient characteristics are related to increased use of services, resources, and costs, we used total episode spending as a composite measure of patient complexity. The use of a composite measure allows us to better understand the relationship between the CJR model and changes in patient characteristics as a whole. We estimated how much of the relative change in total payments experienced over the intervention period was attributable to relative changes in patient mix using a Kitagawa-Blinder-Oaxaca decomposition method. Kröger and Hartmann (2021) developed and described the approach in detail.[31]

First, we conducted an OLS regression to estimate the relative difference in total payments between mandatory CJR and control hospitals over the intervention period, risk adjusting for hospital- and market-level covariates. This model did not risk adjust for patient-level covariates, as those variables were used later to analyze the impact of changes in patient mix. As such, the estimated relative difference captured both the impact of the model on total payments, similar to that captured in our total payments DiD approach, as well as the impact of any effects resulting in relative changes in patient characteristics across CJR and control hospitals.

Next, we decomposed the relative change in total payments from the baseline to the intervention period into separate impacts, each of which are a different type of effect that contributed to the overall relative changes in total payments. The impact of interest for this analysis was a bundle of patient characteristics.[32] To aid in comparison to the DiD, the patient characteristics examined in the decomposition were selected to resemble those used in impact analyses. It informed us of the degree to which the relative change in total payments resulted from relative changes in patient characteristics. The method allowed us to "turn off" any changes between CJR and control hospitals that were not due to changes in patient mix and isolate the effect of changes in patient mix on relative changes in total payments. The estimated impact of changes in patient mix should be interpreted in per-episode units of total episode spending. Standard errors were obtained through a bootstrap process with 500 replications.

We next completed the same analysis, but instead of estimating the impact of changes in patient mix in the entirety of the intervention period relative to the baseline period, we estimated separate effects for each performance year. This approach is also described in Kröger and Hartmann (2021). This allows us to understand how changes in patient characteristics influencing total payments has evolved over the life of the CJR model.

---

[31] Kröger, Hannes, and Jörg Hartmann. 2021. "Extending the Kitagawa-Oaxaca-Blinder decomposition approach to panel data." The Stata Journal. June 29. doi.org/10.1177/1536867X211025800.

[32] Specifically, the following patient characteristics: HCC score, age, sex, race and ethnicity, disability status at Medicare enrollment (not ESRD), Medicaid eligibility status, obesity, hypertension, tobacco use, a flag for whether the LEJR was a knee replacement, and prior utilization measures.

The Oaxaca decomposition approach followed the empirical strategy of the standard DiD approach, but in our application had a few notable differences. In our DiD analyses, we included patient characteristics as risk-adjusting covariates, which caused relative changes in patient characteristics to not influence the DiD impact estimate. Alternatively, the Oaxaca decomposition analyses isolated the effects of relative changes in patient characteristics. Any effect found represented entirely separate ways in which the CJR model impacted total payments, and thus should not be considered part of the DiD impact estimate. Likewise, as the Oaxaca decompositions found the effects of relative changes in patient characteristics on relative changes in total payments, independent of other changes, the DiD impact estimate should not be considered part of the estimates produced by the Oaxaca decompositions. Instead, each estimate contributed additional and unique information, and when taken together, provided a more complete picture of the various ways in which the CJR model influenced relative changes in total payments.

Because hospitals have different quality-adjusted target prices by fracture status and MS-DRG, we performed this analysis separately for four episode groups based on fracture status and presence or absence of a major complication or comorbidity. To gain a more complete understanding of the relationship between the CJR model and changes in patient mix, we ran our analyses using only inpatient LEJR episodes included in the CJR model. Due to limitations in sample size, we only estimated PY specific effects for the sample of elective episodes without a major complication or comorbidity (elective MS-DRG 470).

We performed a sensitivity analysis accounting for the differential outpatient TKA and THA rates, because outpatient TKA and THA patterns can affect the composition of the inpatient LEJR patient population. Accounting for the differential outpatient TKA and THA rates, changes in the patient population resulted in a \$161 ($p<0.01$) relative decrease in average CJR episode payments for elective MS-DRG 470 episodes. In other words, the differential outpatient TKA and THA rates explained about \$32 of the \$193 ($p<0.01$) per-episode decrease due to changes in patient mix for inpatient elective MS-DRG 470 episodes.

B. Analyses of patient characteristics

For our univariate analysis, we estimated DiD regressions[33] on various beneficiary characteristics. We used logistic, multinomial logit, and OLS regression models, chosen to account for the data properties of each characteristic. This analysis did not contain risk-adjusting covariates, since the dependent variables are beneficiary characteristics that are included in our risk adjustment. Standard errors were allowed to cluster at the MSA level.

---

[33] See Section III.3 for additional details about our DiD design.

VI. Impact of the CJR Model on Patients with LEJRs due to Hip Fracture

A. Description of patients with LEJRs due to hip fracture and their care pathways

We compared characteristics of fracture LEJR to elective LEJR that met episode eligibility for the mandatory CJR group at baseline. Characteristics included demographic (age, sex, dual eligibility, disability), health status (HCC score, diagnosis of cancer, congestive heart failure, dementia, diabetes, hypertension or obesity), prior health care utilization (prior use of IRF, NF, SNF, HH, inpatient acute care hospital, and any prior use), and length of anchor hospitalization (in days). We focused on the baseline period so we could look at the fracture population before the CJR model impacted their utilization, care pathways, and recovery. We reported summary statistics (percentages and means). Differences were tested using t-tests for continuous variables and chi-square tests for categorical variables and considered statistically significant if the p-value was less than 0.10.

We studied care pathways for fracture patients at mandatory CJR hospitals from their hospital discharge up to the following 360 days post-discharge. A "care pathway" was defined as unique sequences of up to 3 stays or periods of care (e.g., IRF, HH, community; HH, community). This analysis included fracture patients that would have been episodes if they had not died during the episode, and we looked separately at care pathways during the baseline and intervention periods. The pathways consisted of unique sequences of different care settings (e.g., NF, IRF, HH, inpatient stay, hospice). We defined a stay or period of care at an IRF, SNF, NF, hospice, LTCH, and other IP setting if the admission or use of the setting occurred within the first five days of the previous setting and no other PAC or setting's care was used in between. We defined a period of care from a HHA if the use of HH occurred within the first 14 days of the previous setting and no other PAC or setting's care was used in between. The days in-between preceding and the consecutive days in the setting were counted as a unique stay. Days that had claims for multiple settings were assigned using the following hierarchy: hospice, inpatient hospital stay, SNF, IRF, LTCH, other IP settings, HH, and NF stays (from MDS assessment data). On days when a patient did not receive care in an institutional setting or HH, we categorized them as "in the community." While lack of institutional or HH care could indicate care was not required, our data precludes us from knowing the health status of patients in the community.

B. Risk-adjusted outcomes specific to patients with LEJRs due to hip fracture

For outcomes during the 90-day episode, we used the DiD approach discussed in Section III.C.1. We ran separate regressions for elective and fracture episodes, excluding those patients not meeting episode eligibility. For elective episodes, we used the same set of risk-adjusting covariates as used for the all-LEJR analyses. For analyses of episodes with hip fracture, we used this same set of risk-adjusting covariates with additional HCC indicators included to tailor the risk-adjustment to the fracture population.

For the mortality outcome and outcomes explored in the post-episode period (zero through 360 days post-discharge) we used the DiD approach discussed in Section III.C.1. with several refinements. First, because fracture patients often face declining health and high mortality rates, we included patients that died during the episode period. Second, we adopted several outcome measures used in other research to further examine the quality, utilization, and mobility of fracture patients. Appendix F has additional details about these new outcomes. Lastly, we studied the impact of the CJR model on the long-term health trajectory of fracture patients by examining the impact of the CJR model during days 91-180, 181-270, and 271-360 after a patient's anchor hospitalization. Because this required nine months of additional data, our post-episode analyses included fracture LEJR procedures with an episode period ending by December 2019.

VII. 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 capture information that is not available from other data sources, such as claims data.

A. Survey sample

We administered the patient survey in two batches to a census of CJR and control patients who had inpatient LEJR surgery for hip fractures. Each batch covered two months of hip fracture episodes (July-August 2021 and September-October 2021). Exhibit C-11 describes the patient survey batches.

Exhibit C-11: Patient survey sample of LEJR patients with hip fractures, by survey batch

Batch	Discharge date	CJR LEJR episodes	CJR patients sampled	Control LEJR episodes	Control patients sampled
1	July or August 2021	883	883	996	996
2	September or October 2021	936	936	888	888

Source: CJR evaluation team analysis of survey data for patients with discharge from LEJR surgery for hip fracture in July-October 2021.

Note: LEJR = lower extremity joint replacement.

1. Survey administration

We mailed surveys to patients between 60 and 120 days after their LEJR discharge (an average of 95 days after discharge). 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 128 days after hospital discharge. Sensitivity analysis did not find any evidence that average time between discharge and survey response differed between the CJR and control patients, nor did we find any evidence that results varied when we controlled for time between discharge and survey receipt.

2. Response rates and analytic samples

The response rate was 34.9% for CJR patients and 37.8% for control patients, not a statistically significant difference (Exhibit C-12). There were 1,819 surveys completed by CJR respondents with episodes during the four months covered by the two survey batches, including patients from 203 of the 274 mandatory CJR participant hospitals with at least one hip fracture episode during our sampling period. There were 1,884 surveys completed by control respondents, including patients from 211 of 267 control hospitals with at least one hip fracture episode during our sampling period.

Exhibit C-12: Sample size and response rate for LEJR patients with hip fractures

Group	Patients surveyed (starting sample)		Survey responses received (analytic sample)		Response rate		
	CJR	Control	CJR	Control	CJR	Control	p-value
Hip fracture	1,819	1,884	634	712	34.9%	37.8%	0.133

Source: CJR evaluation team analysis of survey data for patients with discharge from LEJR surgery for hip fractures in July-October 2021.

Note: LEJR = lower extremity joint replacement.

B. Analytic methodology

This section describes the general analytic approach for the survey of LEJR patients with fractures.

1. 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), described in Exhibit C-13. The patient survey instrument is available in Appendix M.

Exhibit C-13: Patient survey domains and measures

Domain	Survey measures[a]	Description of survey measures
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/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
Overall recovery	Satisfaction with overall recovery	Satisfaction with overall recovery since leaving the hospital
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
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
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 refer directly to the joint that received surgery. All other items refer 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/pain at the time of the survey and their recalled level of function/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 is 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 (e.g., "very dissatisfied" or "complete help needed") yielded a score of 0, and the highest response category (e.g., "very satisfied" or "no help needed") yielded a score of 100.

There were three measures of care transition. The first measure, timing of discharge, included 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 two weeks after surgery was more than respondent needed, about right, or not enough). The third measure, did the respondent have all the medical equipment he/she needed when sent home, had two response options (yes or no).

2. Composite measures

We created composite measures for two domains. Reliance on caregiver help, conditional on having any caregiver help, summarized responses to three questions. Satisfaction with care management summarized responses to four questions. To create the composite measure, we translated response items into numeric scores and set them so that zero represented "the most negative amount of the construct being measured" (e.g., 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 ranges from zero (maximum help needed) to six (no help needed for any of the three tasks). Consistent with the Consumer Assessment of Health Providers and Systems (CAHPS) scoring, we re-scaled the composite items so that scores ranged from zero to 100, where zero indicated the least

favorable outcome of the construct being measured (i.e., greatest reliance on caregiver help, and least satisfaction with care management).

Exploratory factor analysis of early returns from the first wave of the survey in PY1 (which comprised approximately 85% of the total wave 1 responses) indicated that the survey items we grouped into composites were internally consistent and, for each composite, reflected a single construct that we could be summarized with one number.

3. 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 minimize differences between the CJR and control patients on key attributes (see domains 1-4 in Exhibit C-14) and minimize differences in observable patient characteristics between CJR or control respondents relative to the full CJR population. For this analysis, we weighted the sample of CJR and control respondents to reflect the CJR population of beneficiaries with hip fractures.

C. Results estimation

For each of the patient survey measures, we estimated the difference between CJR and control patients who received LEJR surgery after a hip fracture.

For our analysis, we utilized the non-linear model listed below for beneficiaries *i*, hospitals *k*, and wave *t* using a general functional form:

$$Y_{i,k,t} = f(b_1 + b_2 \cdot CJR_i + X_{i,k,t}' \cdot B + u_{i,k,t}) \tag{1}$$

Where:

Coefficient b_2 captures the difference in outcomes between CJR and control episodes,

$X_{i,k,t}$ indicates risk factors controlled for in our model, and *f* is a non-linear function.

D. 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 C-14). We selected these 12 patient-level characteristics as covariates for all survey analyses, based on the factors most strongly correlated with patient experience on the prior Bundled Payments for Care Improvement initiative patient survey and conceptual considerations (i.e., factors predicted to be important based on theory).

Exhibit C-14: Risk adjustment to control for differences in patient demographics and clinical risk factors

Domain	Variables
Characteristics of the procedure	MS-DRG[a]
Patient demographics and Medicare enrollment status	Age Sex Dual Medicare/Medicaid eligibility Originally qualified for Medicare due to disability Assignment to ACO Self-reported race and ethnicity[b] Self-reported education[b] Self-reported pre-hospital functional status[b]
Prior health conditions	HCC score Stay in skilled nursing facility or nursing home in six months prior to admission
Survey dimensions	Proxy status (patient had help from someone else in responding to the survey)
Optional patient, hospital, and MSA-level covariates[c]	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 BPCI Advanced 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, 2021 POS, Fiscal Year 2016 CMS Annual IPPS, CMS Master Data Management, 2015-2016 Area Health Resource Files, and Bundled Payments for Care Improvement and Bundled Payments for Care Improvement Advanced initiative participant lists.

Notes: ACO = accountable care organization, BPCI = Bundled Payment for Care Improvement initiative, HCC = hierarchical condition category, IPPS = Inpatient Prospective Payment System, LEJR = lower extremity joint replacement, MSA = metropolitan statistical area, MS-DRG = Medicare Severity-Diagnosis Related Group, PGP = physician group practice, POS = Provider fo Services

[a]   MS-DRG 521 (LEJR due to hip fracture with MCC) and 522 (LEJR due to hip fracture without MCC) took effect in October 2020 and were used for risk adjustment in this analysis.

[b]   For risk adjustment measures that are self-reported (i.e., pre-hospital functional status; race/ethnicity; education), we coded all missing responses as 0 and included an additional binary variable indicating "missing item" (e.g., missing race and ethnicity).

[c]   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 these 12 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 12 fixed variables, as well as the optional patient, hospital, and MSA-level variables listed in the fifth domain in Exhibit C-14.

E. 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 pre-hospital functional status, our results may still be biased if CJR and control respondents had substantially different functional status prior to surgery. For each of the eight pre-hospital functional status measures, we calculated the standardized difference in the unweighted mean between CJR and control respondents. Standardized differences for pre-hospital functional status between CJR and control respondents were all close to or below 0.10, which is a conservative threshold for identifying potentially problematic differences between two groups (Exhibit C-15).[34]

Exhibit C-15: Summary statistics in pre-hospital functional status between CJR and control respondents with hip fractures

Measure	Response rangeᵃ	Hip fracture		
		CJR mean	Control mean	Std. diff.
Walking without rest	-4 to 4	3.14	3.15	0.09
Going up or down stairs	-3 to 3	2.85	2.86	0.09
Rising from sitting	-4 to 4	3.95	3.99	0.04
Standing	-4 to 4	4.04	4.10	0.11
Use of a mobility device	-2 to 2	2.31	2.34	0.04
Getting on or off the toilet	-4 to 4	4.10	4.14	0.07
Pain limiting regular activities	-4 to 4	4.27	4.28	0.12
Medication intensity	-3 to 3	3.66	3.67	0.05

Source: CJR evaluation team analysis of patient survey data for LEJR episodes with discharge in July-October 2021.

Notes: Means and standardized differences are unweighted.

LEJR = lower extremity joint replacement.

ᵃ Difference between a respondent's self-reported status at the time of the survey and the respondent's recalled status prior to the hospitalization.

F. 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 C-16). For each of these measures, we calculated the standardized difference in the

---

[34] Austin, P. C. 2011. "An Introduction to Propensity Score Methods for Reducing the Effects of Confounding in Observational Studies." *Multivariate Behav Res* 46(3): 399-424.

unweighted mean between CJR and control respondents. Standardized differences were below 0.10 for all variables.

Exhibit C-16: Summary statistics in claims-based patient characteristics between CJR and control respondents with hip fractures

Measure	Hip fractures		
	CJR mean	Control mean	Std. diff.
MS-DRG 521 (with MCC)	0.15	0.18	0.09
Age	80.81	80.86	0.01
Female	0.69	0.69	0.01
Eligible for Medicaid	0.07	0.06	0.05
Disability, no ESRD	0.08	0.07	0.04
Assignment to ACO	0.43	0.47	0.08
HCC score	2.42	2.41	0.00
Prior SNF stay[a]	0.07	0.06	0.04

Source: CJR evaluation team analysis of patient survey data for LEJR episodes with discharge in July-October 2021.

Notes: Means and standardized differences are unweighted.

ACO = accountable care organization, ESRD = end-stage renal disease, HCC = hierarchical condition category, LEJR = lower extremity joint replacement, MS-DRG = Medicare Severity-Diagnosis Related Group, SNF = skilled nursing facility.

[a] Stay in skilled nursing facility or nursing home in six months prior to admission.

G. Sensitivity analysis

Unlike CJR hospitals, CMS allowed control hospitals to enroll in Bundled Payments for Care Improvement 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 Bundled Payments for Care Improvement Advanced LEJR. Hospitals participating in the Bundled Payments for Care Improvement Advanced LEJR clinical episode initiated 11.0% of all control episodes, and 6.0% of control episodes were attributed to physician group practices participating in the Bundled Payments for Care Improvement Advanced LEJR clinical episode. As a sensitivity analysis, we excluded all of these episodes, which did not change our results (see Appendix N, Exhibits N-2). This suggests that Bundled Payments for Care Improvement Advanced involvement within control hospitals did not bias our CJR patient survey results.

H. Limitations

The analyses have potential limitations related to the sample, timing of the survey, potential for recall bias, and differential characteristics of CJR and control respondents. Roughly two in three patients did not respond to the survey. 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 hip fracture patients in the CJR model. Since most survey measures focus on past events (e.g., recalled functional status a week prior to surgery, PAC received weeks or months prior to the survey), 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 applies to both CJR and control groups, but it would reduce the precision of the estimates (greater confidence intervals).

VIII. Impact of the CJR Model on Historically Underserved Populations

A. Representativeness of Historically Underserved Populations in the CJR Model

We studied the representation of historically underserved populations in eight unique samples. In each sample, we calculated the proportion that was: 1) dually-eligible for Medicaid; 2) Black or African American; and 3) both dually-eligible for Medicaid and Black or African American. The samples, labeled A through H, are defined below.

Sample A: A beneficiary-year level sample composed of all Medicare FFS beneficiaries between 2016 and 2020 who were alive at any point in the year.

Sample B: An LEJR-level sample composed of all LEJRs received by Medicare FFS beneficiaries that were initiated on or after April 1, 2016 and with a 90-day post-discharge period that ended by December 31, 2020.

Sample C: The sample described in Sample B with the additional restriction that the LEJR met CJR model eligibility rules. LEJRs that met any of the following were considered ineligible under CJR model rules and were thus excluded from this sample: death occurred during the anchor hospitalization or 90-day post-discharge period; LEJR was a BPCI episode; beneficiary was claimed by an ACO beginning July 2017; beneficiary was eligible for Medicare due to ESRD; beneficiary had a subsequent CJR readmission within the 90-day post-discharge period; or beneficiary was transferred to another hospital.

Sample D: The sample described in Sample C with the additional restriction that the LEJR was performed in the inpatient setting. This removed outpatient TKAs and THAs.

Sample E: The sample described in Sample D with the additional restriction that the hospital in which the LEJR was performed was located in one of the original 171 MSAs (67 CJR MSAs and 104 control group MSAs).

Sample F: The sample described in Sample E with the additional restriction that the hospital in which the LEJR was performed was a participant in the CJR Model. This included hospitals in mandatory CJR MSAs, opt-in CJR hospitals in voluntary MSAs, and non-opt-in hospitals in voluntary MSAs before they left the model.

Sample G: The sample described in Sample F with the additional restriction that the hospital in which the LEJR was performed as a mandatory participant in a CJR MSA. This excluded LEJRs from opt-in low volume or rural hospitals in mandatory CJR MSAs.

Sample H: The sample described in Sample G with the additional restriction the LEJR was in the evaluation sample used for impact analyses. This excluded LEJRs for which we could not construct outcomes or risk-adjustment variables (e.g., beneficiaries who received LEJR within six months of joining Medicare). To match our impact analyses, shares of historically underserved populations in this sample were obtained using the sampling weights described in Section II.B.

The shares of each historically underrepresented group in each sample are displayed in Exhibit C-17.

Exhibit C-17: **Description of the samples and population representation**

Sample	A	B	C	D	E	F	G	H
Medicare FFS beneficiaries	✓	✓	✓	✓	✓	✓	✓	✓
+ LEJR		✓	✓	✓	✓	✓	✓	✓
+ Meet model eligibility			✓	✓	✓	✓	✓	✓
+ Inpatient				✓	✓	✓	✓	✓
+ In voluntary or mandatory CJR MSAs					✓	✓	✓	✓
+ CJR episodes						✓	✓	✓
+ In mandatory CJR MSAs							✓	✓
+ In evaluation sample								✓

Frequency in sample	A	B	C	D	E	F	G	H
Total sample	199,456,061	2,486,905	1,632,781	1,316,905	862,509	344,325	205,062	187,477
Black or African American	9.3%	5.4%	4.9%	5.1%	5.3%	5.2%	5.8%	5.7%
Eligible for Medicaid	17.8%	10.1%	10.6%	11.6%	10.5%	11.1%	11.9%	11.5%
Black of African American and eligible for Medicaid	3.2%	1.6%	1.6%	1.7%	1.7%	1.7%	1.9%	1.8%

Source: Master Beneficiary Summary File for the years 2016 to 2020 and evaluation analysis of Medicare claims and enrollment data for LEJRs initiated after April 1, 2016 and with a 90-day post-discharge period that ended by December 31, 2020.

Notes: FFS = fee for service, LEJR = lower extremity joint replacement, MS-DRG = Medicare Severity Diagnosis Related Group, MSA = metropolitan statistical area.

B. Analysis of the Impact of CJR Model on Claims-based Outcomes for Historically Underserved Populations

Our analysis of the differential impact of the CJR model on subpopulations with historically poorer access to care and health outcomes is based on the DiD methodology. We studied the differential impact of the CJR model by estimating the impact of the CJR model on a historically underserved subpopulation and a reference subpopulation, and then estimated the difference between the two CJR model impacts to determine if the CJR model impacted the historically underserved subpopulation differently than 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.

For this report, we studied three historically underserved subpopulations: 1) patients who were Black or African American, 2) patients who were eligible for both Medicare and Medicaid (dually

eligible), and 3) patients who were Black or African American *and* dually eligible. The reference subpopulations were, respectively: 1) White patients, 2) patients who were not dually eligible, and 3) patients who were both White and not dually eligible. The estimated differential impact represents how the difference in the risk-adjusted average outcome between the historically underserved subpopulation (e.g., Black or African American patients) and reference subpopulation (e.g., White patients) changed between the baseline and intervention periods due to the CJR model. In other words, it represented the difference between the effect of the CJR model on the historically underserved subpopulation and the effect of the CJR model on the reference subpopulation, also known as a triple difference or 'DDD' analysis.

In general, the statistical model to estimate the triple difference is:

$$Y_{et} = b_0 + b_1\boldsymbol{B}_{et} + b_2\boldsymbol{H}_{et} + b_3(G_{et} \times CJR_e \times After_t) + \varepsilon_{et}$$

Where Y_{et} is the outcome of interest for episode e at time t. $\boldsymbol{B}_{et}$ is a vector of episode-level risk-adjustment characteristics, including beneficiary demographics, HCC and COVID-19 flags, and measures of prior utilization. $\boldsymbol{H}_{et}$ is a vector of MSA-level risk-adjustment characteristics, including state and region fixed effects, indicators for prior participation in BPCI-A, geographic COVID-19 case rates, and hospital-level flags for ownership and size categories.

The variable G_{et} took the value of 1 if the beneficiary for episode e belongs to the historically underserved subgroup and the value of 0 if the beneficiary belonged to the reference subgroup. For instance, in the analysis of dually eligible beneficiaries compared to beneficiaries without dual eligibility, dually eligible beneficiaries would have $G_{et} = 1$ and beneficiaries without dual eligibility would have $G_{et} = 0$. The coefficient b_3 would capture the difference between the estimated effects of the two subgroups, or the differential impact of the model.

We explored differential impacts of the CJR model on eight outcomes. Total payments, HH visits, SNF length-of-stay, and IRF length-of-stay used ordinary least squares regressions. Mortality, ED use, and unplanned readmissions used logistic regression. Finally, first PAC discharge destination used a multinomial logistic regression.

In the analysis of the impact of CJR on subpopulations with historically poorer access to care and health outcomes, we controlled for changes in patient characteristics. However, despite rigorous risk adjustment, a change in the complexity of the subpopulation could still affect the DiD estimates. For example, DiD estimates that showed a decrease in payment or an increase in quality for a subpopulation could be due in part to CJR hospital participants selecting less complex patients from this subpopulation in the intervention period.

C. Analysis of the Impact of CJR Model on LEJR Volume for Historically Underserved Populations

Prior research reported widening disparities in access to LEJR for certain historically underserved populations under the CJR model.[35,36] We evaluated the impact of the CJR model on volume of LEJR discharges for three historically underserved populations.

1. Sample and Time Periods

Our sample included all Medicare FFS beneficiaries who resided in a mandatory CJR or control MSA between 2012 and 2019. Because the CJR model's first performance period (PY1) began in the middle of a calendar year, our analyses using the beneficiary-year sample excluded 2016 data.

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 61,372,896 beneficiary-years.

2. LEJRs per 100,000 FFS population

Results are presented 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.

3. Statistical model (Overall DiD)

The outcome of interest was the estimated probability that a beneficiary received at least one 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.

$$Logit(Y_{it}) = b_0 + b_1C_{it} + b_2HCC_{it} + b_3PC_{it} + b_4(CJR_{it} \times After_t) + b_5S_{it} + \varepsilon_{it}$$

Where Y_{it} is an indicator variable that takes on the value of 1 if beneficiary i received at least one elective LEJR in year t. C_{it} is a vector of beneficiary characteristics including indicator variables for race, dual eligibility for Medicare and Medicaid, age buckets, sex, and certain chronic complications like obesity, diabetes, or tobacco use. HCC_{it} is a vector of indicator variables for a subset of HCC flags. PC_{it} is a vector of variables for prior care use – specifically, continuous variables reflecting the number of days receiving health care services in the six months prior to the LEJR. S_{it} is a vector of binary variables indicating the state in which the beneficiary resided.

[35]   Caroline Thirukumaran, Yeunkyung Kim, Xueya Cai, Benjamin Ricciardi, Yue Li, Kevin Fiscella, Addisu Mesfin, and Laurent Glance. *Association of the Comprehensive Care for Joint Replacement Model With Disparities in the Use of Total Hip and Total Knee Replacement*. JAMA Netw Open. 2021;4(5):e2111858. doi:10.1001/jamanetworkopen.2021.11858

[36]   Hyunjee Kim, Thomas Meath, Ana Quiñones, John McConnell, and Said Ibrahim. *Association of Medicare Mandatory Bundled Payment Program With the Receipt of Elective Hip and Knee Replacement in White, Black, and Hispanic Beneficiaries*. JAMA Netw Open. 2021;4(3):e211772. doi:10.1001/jamanetworkopen.2021.1772

The impact of the CJR model was captured by the coefficient b_4, which was identified by comparing the beneficiaries who resided in CJR MSAs during the intervention period ($After_t = 1$) 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. The coefficient was then transformed into an LEJR rate (number of LEJR performed per 100,000 FFS beneficiaries per year). Standard errors were clustered at the MSA level.

4. *Statistical model (Health Equity DDD)*

The triple difference analysis is similar to performing DiD on two subpopulations and then comparing the resulting estimates.[37]

The model specification for the health equity triple-difference analyses is:

$$logit(Y_{it}) = b_0 + b_1 X_{it} + b_2(G_{it} \times CJR_{it} \times After_t) + \varepsilon_{it}$$

Where Y_{it} is an indicator variable that takes on the value of 1 if beneficiary i received at least one elective LEJR in year t and X_{it} is a vector of control covariates (containing all of the covariates mentioned in Section C, above). G_{it} is a binary variable that indicates whether beneficiary i in year t is a member of the 'historically underserved subpopulation' or the 'reference subpopulation'. The coefficient b_2 captures the difference between the estimated effect of the CJR model on the historically 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_2 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.

5. *Limitations*

An important limitation to the beneficiary-year analyses is that the sample, by construction, does not capture LEJRs performed on beneficiaries who do *not* reside in a mandatory CJR or control MSA. A non-negligible portion of LEJRs performed in mandatory CJR or mandatory control MSAs involved beneficiaries travelling from other locations, and these LEJRs were 'missed' by the beneficiary-year analysis. Fully capturing these LEJRs in a beneficiary-year file would require constructing a beneficiary-year sample of *all* Medicare FFS beneficiaries regardless of residence.

---

[37] In precise terms, our triple-difference analysis did not equate to running two separate DiD analyses, because we did not include an interaction of the group indicator (G_{it}) with the control covariates (X_{it}). Running two separate DiDs implicitly allowed for such an interaction.

IX. Qualitative and Mixed Methods Analysis

A. Analysis of provider telephone 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 (i.e., codes were dropped, consolidated, added, or revised) to better capture patterns as they emerged.

1. Limitations

The analysis of the telephone interview data provides a description of themes and patterns in response to the protocols, which may not include the full experience of CJR hospital participants. Our sample of hospitals was limited to 39 total interviews and may not be representative of CJR participating hospitals that cared for patients with LEJR due to hip fracture.

B. Analysis of the care coordination survey

Survey data were reviewed and cleaned before analysis. For hospitals with multiple responses, we retained the response without missing values. If multiple responses for an individual had missing values, then we retained the earliest response submitted.

For survey measures, we calculated frequencies and summary statistics for all close-ended questions included in the survey. For open-text items, we reviewed responses and identified common themes.

1. Non-response analysis

The response rate for the care coordination survey was 50.4%. A non-response analysis was conducted to assess the generalizability of the care coordination survey respondents to the overall hospital survey sample. We compared characteristics of survey respondents (n=199) and non-respondents (n=196) to determine if there was the potential for response bias that would need to be accounted for in our analysis. Statistical significance of characteristics associated with survey response was determined using Pearson's chi-squared tests.

Characteristics of respondents were proportionally similar to non-respondents regarding prior participation in Bundled Payments for Care Improvement initiative for LEJR (15% and 14%, respectively), receipt of reconciliation payments in PY4 (59% and 53%), opt-in hospital status (23% and 20%), location in a high payment MSA (81% and 86%), and health system affiliation (93% and 88%). The only statistically significant difference was related to LEJR volume. Fifty-four percent of respondents performed more than 106 LEJRs in PY4, compared to 45% of non-respondents ($p<0.10$).

Given the similarly of the two populations, we believed respondents and non-respondents did not vary in meaningful ways and creation and use of a non-response weight was not warranted.

2. Limitations

The response rate of the care coordination survey was 50%. Results may not be generalizable to all hospitals participating in the CJR model. While we found that responding care coordinators were similar to non-respondents on most characteristics, respondents may differ from non-respondents in unobserved ways that we could not measure and explore.

We had to rely on hospital points of care to provide names and contact information for their care coordinators. This additional step to identify survey respondents may have impacted our response rate.

There is the possibility of recall bias, as survey respondents were asked to reflect on their past experience and the influence of the CJR model on the adoption and enhancement of care coordination activities when providing answers to survey questions. Social desirability bias and recency bias may lead responding care coordinators to provide more positive reports than their actual experience or to forget changes or experiences from the early years of the model. Further, respondent length of time in the care coordinator position (e.g., a recent hire vs. hired prior to or at the start of the CJR model) may have impacted his/her ability to recall specific details about care coordination activities and the influence of the CJR model.

Lastly, the study design was a cross-sectional survey and thus analyses conducted for the care coordinator survey cannot inform statements about cause and effect.

Appendix D: Payment, Utilization, and Quality Results

Performance-Year Specific Results

Exhibit D-1: Risk-adjusted claims-based difference-in-differences results for payment, utilization, and quality metrics, mandatory CJR hospitals, all LEJR episodes, by performance year

Domain	Measure	Baseline risk-adjusted average		PY1		PY2		PY3		PY4		PY5.1		PY5.2	
		CJR	Control group	DiD	P-value	DiD	P-value	DiD	P-value	DiD	P-value	DiD	P-value	DiD	P-value
Payments	Episode payments	$29,556	$29,008	-$1,404	<0.01	-$1,594	<0.01	-$1,324	<0.01	-$1,194	<0.01	-$1,092	0.03	-$973	0.08
	IRF payments	$2,360	$2,240	-$595	<0.01	-$541	<0.01	-$553	0.02	-$753	<0.01	-$573	0.07	-$427	0.27
	SNF payments	$6,339	$6,414	-$711	<0.01	-$973	<0.01	-$820	<0.01	-$460	0.09	-$282	0.36	-$325	0.32
	HH payments[a]	$2,426	$2,329	$124	0.16	$41	0.71	$84	0.57	$40	0.81	$93	0.59	$231	0.19
	Readmission payments	$1,264	$1,133	-$118	0.11	-$158	0.07	-$159	0.08	-$167	0.04	-$102	0.29	-$137	0.17
	Part B payments	$5,057	$4,872	-$158	<0.01	-$101	0.11	-$30	0.62	$28	0.69	-$94	0.27	-$37	0.71
	30-day PEP payments[a]	$1,538	$1,540	$7	0.84	-$42	0.12	-$62	0.09	-$51	0.09	-$33	0.35	$54	0.50
	Anchor payments	$12,219	$12,223	$11	0.85	$34	0.52	$101	0.03	$163	<0.01	$60	0.06	$42	0.33
Utilization	First PAC IRF	14.4%	13.5%	-3.7	<0.01	-3.6	<0.01	-3.5	0.03	-4.8	<0.01	-3.4	0.08	-2.7	0.21
	First PAC SNF	41.2%	42.2%	-2.1	0.06	-3.3	<0.01	-2.5	0.17	1.2	0.56	-0.7	0.76	-2.6	0.25
	First PAC HH	36.2%	33.2%	7.7	<0.01	7.5	0.01	7.7	0.05	4.3	0.33	6.7	0.19	11.3	0.03
	First PAC home without HH	8.2%	11.2%	-1.9	0.30	-0.6	0.77	-1.7	0.57	-0.6	0.83	-2.6	0.43	-6.0	0.07
	Any HH use	74.5%	72.0%	5.1	0.03	4.2	0.13	4.8	0.27	2.9	0.56	5.0	0.35	9.4	0.09
	IRF days	11.9	11.7	-0.1	0.51	-0.1	0.75	-0.2	0.36	-0.1	0.79	-0.4	0.18	-0.2	0.37
	SNF days	27.8	27.9	-2.0	0.01	-2.6	<0.01	-2.6	<0.01	-2.6	<0.01	-1.3	0.05	-0.4	0.63
	HH visits	17.1	16.9	-0.6	0.06	-1.2	0.01	-1.1	0.02	-0.9	0.07	-0.7	0.15	-0.5	0.37
	Outpatient PT/OT visits[a]	13.0	13.2	0.3	0.16	0.5	0.14	0.4	0.21	0.4	0.19	0.3	0.18	0.1	0.70
Quality	Unplanned readmission rate	9.4%	9.1%	-0.3	0.21	-0.4	0.15	-0.1	0.64	-0.2	0.40	0.2	0.45	0.1	0.84
	ED use	13.2%	12.8%	-0.1	0.89	-0.3	0.47	0.5	0.07	0.5	0.14	0.0	0.91	-0.1	0.88
	Mortality rate	2.8%	2.8%	0.1	0.82	0.0	0.85	0.1	0.49	0.0	0.80	0.1	0.71	-0.2	0.35

Source: CJR evaluation team analysis of Medicare claims and enrollment data for episodes initiated in 2012 through 2014 that ended between April 2012 and March 2015 (baseline) and episodes initiated during or after April 2016 that ended by September 2021 (intervention).

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

The change in separate provider payments do not sum to the change in episode payments because separate models were estimated for episode payments and each component payment.

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

DiD = difference-in-differences, ED = emergency department, HH = home health, IRF = inpatient rehabilitation facility, LEJR = lower extremity joint replacement, OT = occupational therapy, PAC = post-acute care, PEP = post-episode period, PT = physical therapy, PY = performance year, SNF = skilled nursing facility, TKA = total knee arthroplasty.

[a] We cannot be certain that there is no impact of the model because this outcome failed parallel trends tests (Appendix K). Parallel trends is an assumption that underlies our methodological approach. Please see Appendix C (Section III.C.1.c) for additional details.

[b] The complications measure only applies to elective episodes.

Cumulative Results

Exhibit D-2: Risk-adjusted claims-based difference-in-differences results for payment, utilization, and quality metrics, mandatory CJR hospitals, all LEJR episodes, PY1-5

Domain	Measure	Intervention episodes (N)		CJR		Control group		DiD	DiD % of baseline	p-value	90% CI
		CJR	Control group	Baseline risk-adjusted average	Intervention risk-adjusted average	Baseline risk-adjusted average	Intervention risk-adjusted average				
Payments	Episode payments	205,893	274,840	$29,485	$26,588	$28,932	$27,471	-$1,437	-4.9%	<0.01	[-$2,069 to -$805]
	IRF payments	205,893	274,840	$2,377	$1,323	$2,255	$1,795	-$594	-25.0%	0.01	[-$985 to -$203]
	SNF payments	205,893	274,840	$6,249	$4,239	$6,317	$5,053	-$747	-11.9%	<0.01	[-$1,111 to -$382]
	HH payments[a]	205,893	274,840	$2,408	$2,435	$2,311	$2,263	$76	3.1%	0.58	[-$147 to $298]
	Readmission payments	205,893	274,840	$1,268	$1,101	$1,137	$1,116	-$146	-11.5%	0.07	[-$280 to -$12]
	Part B payments	205,893	274,840	$5,026	$4,892	$4,841	$4,785	-$77	-1.5%	0.20	[-$176 to $23]
	30-day PEP payments[a]	205,893	274,840	$1,534	$1,556	$1,536	$1,582	-$25	-1.6%	0.32	[-$66 to $17]
	Anchor payments	205,893	274,840	$12,247	$12,248	$12,251	$12,200	$52	0.4%	0.18	[-$12 to $116]
Utilization	First PAC IRF	205,892	274,840	14.3%	6.2%	13.4%	9.1%	-3.8	-26.7%	0.02	[-6.4 to -1.2]
	First PAC SNF	205,892	274,840	40.8%	28.2%	41.6%	31.6%	-2.5	-6.2%	0.12	[-5.2 to 0.2]
	First PAC HH	205,892	274,840	36.6%	50.0%	33.6%	39.5%	7.5	20.5%	0.04	[1.4 to 13.6]
	First PAC home without HH	205,892	274,840	8.4%	15.6%	11.4%	19.9%	-1.2	-14.4%	0.62	[-5.2 to 2.8]
	Any HH use	205,893	274,840	74.5%	75.5%	71.9%	68.3%	4.6	6.2%	0.24	[-1.9 to 11.2]
	IRF days	14,723	20,479	11.9	12.0	11.7	12.0	-0.2	-1.3%	0.41	[-0.5 to 0.2]
	SNF days	61,263	63,540	27.7	22.8	27.8	25.2	-2.3	-8.2%	<0.01	[-3.0 to -1.5]
	HH visits	150,242	178,987	17.0	15.6	16.8	16.3	-1.0	-5.8%	0.02	[-1.7 to -0.3]
	Outpatient PT/OT visits[a]	137,703	201,174	13.0	14.6	13.1	14.4	0.4	3.1%	0.13	[-0.0 to 0.8]

| Domain | Measure | Intervention episodes (N) | | CJR | | Control group | | | | | |
		CJR	Control group	Baseline risk-adjusted average	Intervention risk-adjusted average	Baseline risk-adjusted average	Intervention risk-adjusted average	DiD	DiD % of baseline	p-value	90% CI
Quality	Unplanned readmission rate	205,823	274,788	9.3%	8.8%	9.1%	8.8%	-0.2	-2.2%	0.27	[-0.5 to 0.1]
	ED use	205,823	274,788	13.1%	13.9%	12.7%	13.3%	0.1	0.7%	0.74	[-0.4 to 0.6]
	Mortality rate	210,998	279,907	2.8%	2.7%	2.9%	2.7%	0.0	0.2%	0.94	[-0.2 to 0.2]

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 initiated during or after April 2016 that ended by September 2021 (intervention).

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

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

The change in separate provider payments do not sum to the change in episode payments because separate models were estimated for episode payments and each component payment.

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

CI = confidence interval, DiD = difference-in-differences, ED = emergency department, HH = home health, IRF = inpatient rehabilitation facility, LEJR = lower extremity joint replacement, OT = occupational therapy, PAC = post-acute care, PEP = post-episode period, PT = physical therapy, PY = performance year, SNF = skilled nursing facility, TKA = total knee arthroplasty.

[a] We cannot be certain that there is no impact of the model because this outcome failed parallel trends tests (Appendix K). Parallel trends is an assumption that underlies our methodological approach. Please see Appendix C (Section III.C.1.c) for additional details.

[b] The complications measure only applies to elective episodes.

Appendix E: MPS Results

Mandatory hospitals

Exhibit E-1: Medicare reductions in payments, reconciliation payments, and program savings by performance year, mandatory hospitals

Year	Component	Estimate	90% Confidence interval
PY1	Reduction in non-standardized paid amounts per episode	$1,366	-$1,922 to -$810
	Reconciliation payments per episode	$872	N/A
	Medicare savings per episode	$494	-$62 to $1,050
	Number of intervention episodes	21,275	N/A
	Aggregate Medicare savings	$10,513,186	-$1,313,707 to $22,340,058
PY2	Reduction in non-standardized paid amounts per episode	$1,548	-$2,089 to -$1,007
	Reconciliation payments per episode	$944	N/A
	Medicare savings per episode	$604	$63 to $1,145
	Number of intervention episodes	43,694	N/A
	Aggregate Medicare savings	$26,374,231	$2,739,084 to $50,009,390
PY3	Reduction in non-standardized paid amounts per episode	$1,246	-$1,875 to -$617
	Reconciliation payments per episode	$703	N/A
	Medicare savings per episode	$543	-$86 to $1,172
	Number of intervention episodes	39,284	N/A
	Aggregate Medicare savings	$21,323,409	-$3,391,713 to $46,038,502
PY4	Reduction in non-standardized paid amounts per episode	$1,038	-$1,752 to -$324
	Reconciliation payments per episode	$756	N/A
	Medicare savings per episode	$282	-$432 to $997
	Number of intervention episodes	48,807	N/A
	Aggregate Medicare savings	$13,771,061	-$21,095,301 to $48,637,413
PY5.1	Reduction in non-standardized paid amounts per episode	$925	-$1,783 to -$67
	Reconciliation payments per episode	$2,716	N/A
	Medicare savings per episode	-$1,791	-$2,649 to -$933
	Number of intervention episodes	34,277	N/A
	Aggregate Medicare savings	-$61,398,647	-$90,811,293 to -$31,986,001
PY5.2	Reduction in non-standardized paid amounts per episode	$834	-$1,818 to $150
	Reconciliation payments per episode	$2,665	N/A
	Medicare savings per episode	-$1,831	-$2,815 to -$847
	Number of intervention episodes	18,556	N/A
	Aggregate Medicare savings	-$33,976,108	-$52,230,756 to -$15,721,461

Source: CJR evaluation team analysis of Medicare claims and enrollment data for episodes initiated in 2012 through 2014 that ended between April 2012 and March 2015 (baseline) and episodes initiated during or after April 2016 that ended by September 2021 (intervention) and CJR payment contractor data for CJR participant hospitals in PY1-5.2.

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

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

DiD = difference-in-differences, N/A = not applicable, PY = performance year, THA = total hip arthroplasty, TKA = total knee arthroplasty.

Exhibit E-2: Cumulative Medicare reductions in payments, reconciliation payments, and program savings, mandatory hospitals

Year	Component	Estimate	90% Confidence interval
PY1-PY5.2	Reduction in non-standardized paid amounts per episode	$1,182	-$1,823 to -$542
	Reconciliation payments per episode	$1,296	N/A
	Medicare savings per episode	-$114	-$754 to $527
	Number of intervention episodes	205,893	N/A
	Aggregate Medicare savings	-$23,392,957	-$155,210,947 to $108,425,095

Source: CJR evaluation team analysis of Medicare claims and enrollment data for episodes initiated in 2012 through 2014 that ended between April 2012 and March 2015 (baseline) and episodes initiated during or after April 2016 that ended by September 2021 (intervention) and CJR payment contractor data for CJR participant hospitals in PY1-5.2.

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

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

DiD = difference-in-differences, N/A = not applicable, PY = performance year, THA = total hip arthroplasty, TKA = total knee arthroplasty.

Opt-in hospitals

Exhibit E-3: Medicare reductions in payments, reconciliation payments, and program savings by performance year, opt-in hospitals

Year	Component	Estimate	Range (90% Confidence interval)
PY1	Reduction in non-standardized paid amounts per episode	$476	-$964 to $12
	Reconciliation payments per episode	$949	N/A
	Medicare savings per episode	-$474	-$962 to $14
	Number of intervention episodes	7,473	N/A
	Aggregate Medicare savings	-$3,539,624	-$7,187,459 to $108,210
PY2	Reduction in non-standardized paid amounts per episode	$773	-$1,119 to -$426
	Reconciliation payments per episode	$1,180	N/A
	Medicare savings per episode	-$407	-$753 to -$60
	Number of intervention episodes	16,891	N/A
	Aggregate Medicare savings	-$6,871,188	-$12,724,300 to -$1,018,073
PY3	Reduction in non-standardized paid amounts per episode	$420	-$842 to $1
	Reconciliation payments per episode	$1,478	N/A
	Medicare savings per episode	-$1,057	-$1,479 to -$636
	Number of intervention episodes	14,425	N/A
	Aggregate Medicare savings	-$15,254,418	-$21,335,536 to -$9,173,301

LEWIN GROUP®

Year	Component	Estimate	Range (90% Confidence interval)
PY4	Reduction in non-standardized paid amounts per episode	$844	-$1,325 to -$364
	Reconciliation payments per episode	$2,362	N/A
	Medicare savings per episode	-$1,517	-$1,998 to -$1,037
	Number of intervention episodes	13,826	N/A
	Aggregate Medicare savings	-$20,979,557	-$27,626,961 to -$14,332,160
PY5.1	Reduction in non-standardized paid amounts per episode	$362	-$876 to $152
	Reconciliation payments per episode	$3,546	N/A
	Medicare savings per episode	-$3,183	-$3,697 to -$2,669
	Number of intervention episodes	8,996	N/A
	Aggregate Medicare savings	-$28,638,425	-$33,262,253 to -$24,014,596
PY5.2	Reduction in non-standardized paid amounts per episode	$676	-$1,494 to $141
	Reconciliation payments per episode	$3,407	N/A
	Medicare savings per episode	-$2,731	-$3,548 to -$1,913
	Number of intervention episodes	3,904	N/A
	Aggregate Medicare savings	-$10,660,588	-$13,852,241 to -$7,468,933

Source: CJR evaluation team analysis of Medicare claims and enrollment data for episodes initiated in 2012 through 2014 that ended between April 2012 and March 2015 (baseline) and episodes initiated during or after April 2016 that ended by September 2021 (intervention) and CJR payment contractor data for CJR participant hospitals in PY1-5.2.

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

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

DiD = difference-in-differences, N/A = not applicable, PY = performance year, THA = total hip arthroplasty, TKA = total knee arthroplasty.

Exhibit E-4: Cumulative Medicare reductions in payments, reconciliation payments, and program savings, opt-in hospitals

Year	Component	Estimate	90% Confidence interval
PY1-PY5.2	Reduction in non-standardized paid amounts per episode	$614	-$977 to -$252
	Reconciliation payments per episode	$1,926	N/A
	Medicare savings per episode	-$1,312	-$1,674 to -$949
	Number of intervention episodes	65,515	N/A
	Aggregate Medicare savings	-$85,943,802	-$109,699,777 to -$62,187,821

Source: CJR evaluation team analysis of Medicare claims and enrollment data for episodes initiated in 2012 through 2014 that ended between April 2012 and March 2015 (baseline) and episodes initiated during or after April 2016 that ended by September 2021 (intervention) and CJR payment contractor data for CJR participant hospitals in PY1-5.2.

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

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

DiD = difference-in-differences, N/A = not applicable, PY = performance year, THA = total hip arthroplasty, TKA = total knee arthroplasty.

Non-opt-in hospitals

Exhibit E-5: Medicare reductions in payments, reconciliation payments, and program savings by performance year, non-opt-in hospitals

Year	Component	Estimate	Range (90% Confidence interval)
PY1	Reduction in non-standardized paid amounts per episode	$70	-$433 to $293
	Reconciliation payments per episode	$497	N/A
	Medicare savings per episode	-$427	-$791 to -$64
	Number of intervention episodes	11,059	N/A
	Aggregate Medicare savings	-$4,726,516	-$8,745,187 to -$707,846
PY2	Reduction in non-standardized paid amounts per episode	$505	-$809 to -$202
	Reconciliation payments per episode	$739	N/A
	Medicare savings per episode	-$234	-$538 to $69
	Number of intervention episodes	23,262	N/A
	Aggregate Medicare savings	-$5,445,932	-$12,504,432 to $1,612,566

Source: CJR evaluation team analysis of Medicare claims and enrollment data for episodes initiated in 2012 through 2014 that ended between April 2012 and March 2015 (baseline) and episodes initiated during or after April 2016 that ended by December 2017 (intervention) and CJR payment contractor data for CJR participant hospitals in PY1-2.

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

DiD = difference-in-differences, N/A = not applicable, PY = performance year.

Exhibit E-6: Cumulative Medicare reductions in payments, reconciliation payments, and program savings, non-opt-in hospitals

Year	Component	Estimate	90% Confidence interval
PY1-PY2	Reduction in non-standardized paid amounts per episode	$365	-$663 to -$67
	Reconciliation payments per episode	$661	N/A
	Medicare savings per episode	-$296	-$594 to $2
	Number of intervention episodes	34,321	N/A
	Aggregate Medicare savings	-$10,170,992	-$20,397,398 to $55,417

Source: CJR evaluation team analysis of Medicare claims and enrollment data for episodes initiated in 2012 through 2014 that ended between April 2012 and March 2015 (baseline) and episodes initiated during or after April 2016 that ended by December 2017 (intervention) and CJR payment contractor data for CJR participant hospitals in PY1-5.2.

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

DiD = difference-in-differences, N/A = not applicable, PY = performance year

All hospitals

Exhibit E-7: Medicare reductions in payments, reconciliation payments, and program savings by performance year, all hospitals

Year	Component	Estimate	90% Confidence interval
PY1	Reduction in non-standardized paid amounts per episode	$839	-$1,186 to -$492
	Reconciliation payments per episode	$782	N/A
	Medicare savings per episode	$56	-$290 to $403
	Number of intervention episodes	39,807	N/A
	Aggregate Medicare savings	$2,248,324	-$11,557,344 to $16,053,993
PY2	Reduction in non-standardized paid amounts per episode	$1,102	-$1,428 to -$776
	Reconciliation payments per episode	$935	N/A
	Medicare savings per episode	$168	-$158 to $493
	Number of intervention episodes	83,847	N/A
	Aggregate Medicare savings	$14,055,282	-$13,266,877 to $41,377,442
PY3	Reduction in non-standardized paid amounts per episode	$1,024	-$1,513 to -$535
	Reconciliation payments per episode	$911	N/A
	Medicare savings per episode	$113	-$376 to $602
	Number of intervention episodes	53,709	N/A
	Aggregate Medicare savings	$6,066,576	-$20,193,751 to $32,326,903
PY4	Reduction in non-standardized paid amounts per episode	$997	-$1,565 to -$428
	Reconciliation payments per episode	$1,110	N/A
	Medicare savings per episode	-$114	-$682 to $455
	Number of intervention episodes	62,633	N/A
	Aggregate Medicare savings	-$7,120,497	-$42,740,250 to $28,499,256
PY5.1	Reduction in non-standardized paid amounts per episode	$820	-$1,513 to -$128
	Reconciliation payments per episode	$2,889	N/A
	Medicare savings per episode	-$2,068	-$2,760 to -$1,376
	Number of intervention episodes	43,273	N/A
	Aggregate Medicare savings	-$89,492,017	-$119,443,225 to -$59,540,808
PY5.2	Reduction in non-standardized paid amounts per episode	$808	-$1,637 to $21
	Reconciliation payments per episode	$2,794	N/A
	Medicare savings per episode	-$1,986	-$2,815 to -$1,157
	Number of intervention episodes	22,460	N/A
	Aggregate Medicare savings	-$44,603,893	-$63,224,390 to -$25,983,396

Source: CJR evaluation team analysis of Medicare claims and enrollment data for episodes initiated in 2012 through 2014 that ended between April 2012 and March 2015 (baseline) and episodes initiated during or after April 2016 that ended by September 2021 (intervention) and CJR payment contractor data for CJR participant hospitals in PY1-5.2.

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

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

DiD = difference-in-differences, N/A = not applicable, PY = performance year, THA = total hip arthroplasty, TKA = total knee arthroplasty.

Exhibit E-8: Cumulative Medicare reductions in payments, reconciliation payments, and program savings, all hospitals

Year	Component	Estimate	90% Confidence interval
PY1-PY5.2	Reduction in non-standardized paid amounts per episode	$971	-$1421 to -$522
	Reconciliation payments per episode	$1,360	N/A
	Medicare savings per episode	-$389	-$838 to $61
	Number of intervention episodes	305,729	N/A
	Aggregate Medicare savings	-$118,846,224	-$256,277,636 to $18,585,188

Source: CJR evaluation team analysis of Medicare claims and enrollment data for episodes initiated in 2012 through 2014 that ended between April 2012 and March 2015 (baseline) and episodes initiated during or after April 2016 that ended by September 2021 (intervention) and CJR payment contractor data for CJR participant hospitals in PY1-5.2.

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

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

DiD = difference-in-differences, N/A = not applicable, PY = performance year, THA = total hip arthroplasty, TKA = total knee arthroplasty.

Appendix F: Outcome Definitions

Exhibit F-1: Claims-based outcome definitions[1]

Domain	Outcome name	Definition	Measurement period(s)	Eligible sample
Medicare payments	Total Medicare standardized allowed amounts per episode[2]	The sum of Medicare payment and beneficiary out-of-pocket amounts for related items and services covered by Medicare Part A and Part B[3] 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 six 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 September 30, 2021; 5) have non-missing Medicare standardized allowed payment information for the episode.
	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 (MS-DRG 469, 470, 521, or 522 for inpatient episodes covered under Medicare Part A; CPT 27447 for outpatient TKA episodes covered under Medicare Part B).	Anchor hospitalization	Beneficiaries who: 1) have a complete FFS enrollment history six 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 September 30, 2021; 5) have non-missing Medicare standardized allowed payment information for the episode.
	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 six 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 September 30, 2021; 5) have non-missing Medicare standardized allowed payment information for the episode.

---

[1] 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.

[2] Standardized payments remove wage adjustments and other Medicare payment adjustments (e.g., GME, IME, and DSH). Allowed amounts include beneficiary cost sharing.

[3] Episode-related items and services paid under Medicare Part A or Part B, after exclusions are applied, include: physician services; inpatient hospital services (including readmissions with certain exceptions discussed in the Final Rule); IPF services; LTCH 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 (cont'd)	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 six 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 September 30, 2021; 5) have non-missing Medicare standardized allowed payment information for the episode.
	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 six 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 September 30, 2021; 5) have non-missing Medicare standardized allowed payment information for the episode.
	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 six 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 October 30, 2021; 5) have non-missing Medicare standardized allowed payment information for the episode.
Utilization	First discharge to IRF	The percent 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 five 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 five days of the transfer discharge.	1st to 5th day after discharge from the anchor/ transfer hospitalization	Beneficiaries who: 1) have a complete FFS enrollment history six 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 September 30, 2021; 5) have non-missing Medicare standardized allowed payment information for the episode.

Domain	Outcome name	Definition	Measurement period(s)	Eligible sample
Utilization (cont'd)	First discharge to SNF	The percent 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 five 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 five days of the transfer discharge.	1st to 5th day after discharge from the anchor/ transfer hospitalization	Beneficiaries who: 1) have a complete FFS enrollment history six 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 September 30, 2021; 5) have non-missing Medicare standardized allowed payment information for the episode.
	First discharge to HHA	The percent of all episodes with beneficiaries initially discharged to an HHA. The first PAC setting is an 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/ transfer hospitalization	Beneficiaries who: 1) have a complete FFS enrollment history six 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 September 30, 2021; 5) have non-missing Medicare standardized allowed payment information for the episode.
	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 an 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/ transfer hospitalization	Beneficiaries who: 1) have a complete FFS enrollment history six 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 September 30, 2021; 5) have non-missing Medicare standardized allowed payment information for the episode.

Domain	Outcome name	Definition	Measurement period(s)	Eligible sample
Utilization (cont'd)	Any HH 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 six 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 September 30, 2021; 5) have non-missing Medicare standardized allowed payment information for the episode.
	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 six 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 September 30, 2021; 5) have at least one IRF day during this period; 6) have non-missing Medicare standardized allowed payment information for the episode.
	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 six 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 September 30, 2021; 5) have at least one SNF day during this period; 6) have non-missing Medicare standardized allowed payment information for the episode.
	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 six 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 September 30, 2021; 5) have at least one HHA visit during this period; 6) have non-missing Medicare standardized allowed payment information for the episode.

Domain	Outcome name	Definition	Measurement period(s)	Eligible sample
Utilization (cont'd)	Number of outpatient PT/OT visits	The average number of outpatient physical therapy and occupational therapy (PT/OT) visits during the 90-day post-discharge period.	90-day post-discharge period	Beneficiaries who: 1) have a complete FFS enrollment history six 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 September 30, 2021; 5) have at least one outpatient PT/OT visit during this period; 6) have non-missing Medicare standardized allowed payment information for the episode.
Quality	Unplanned readmission rate	The proportion of episodes with one or more unplanned readmissions for any eligible condition. This measure was based on specifications for the NQF-endorsed all-cause unplanned readmission measure (NQF measure 1789).[4] Following these specifications, we excluded planned admissions, based on AHRQ Clinical Classification System Procedure and Diagnoses codes.	90-day post-discharge period	Beneficiaries who: 1) have a complete FFS enrollment history six 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 September 30, 2021; 5) are discharged from the anchor hospitalization in accordance with medical advice; 6) have non-missing Medicare standardized allowed payment information for the episode.
	Emergency department 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 emergency department but was not admitted to the hospital.	90-day post-discharge period	Beneficiaries who: 1) have a complete FFS enrollment history six 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 September 30, 2021; 5) are discharged from the anchor hospitalization in accordance with medical advice; 6) have non-missing Medicare standardized allowed payment information for the episode.

---

[4] 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

LEWINGROUP®

Domain	Outcome name	Definition	Measurement period(s)	Eligible sample
Quality (cont'd)	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 six 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 six months prior to admission; 5) have a measurement period that ends on or before September 30, 2021; 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: AMI, pneumonia, or sepsis/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/TKA complications measure (NQF measure 1550).[5] 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 an elective procedure (non-fracture); 2) have a complete FFS enrollment history six months prior to the anchor hospitalization; 3) have consistent, reliable sex and age data (age <115); 4) maintain Parts A and B enrollment throughout the measurement period; 5) have a measurement period that ends on or before September 30, 2021; 6) are discharged from the anchor hospitalization in accordance with medical advice; 7) have non-missing Medicare standardized allowed payment information for the episode.

Note: ACH = acute care hospital, AHRQ = Agency for Healthcare Research and Quality, AMI = acute myocardial infarction, CPT = current procedural terminology, DME = durable medical equipment, DSH = disproportionate share hospital, ED = emergency department, FFS = fee-for-service, GME = graduate medical education, HH = home health, HHA = home health agency, IME = indirect medical education, IPF = inpatient psychiatric facility, IRF = inpatient rehabilitation facility,

---

[5] Updated specification documents were released by CMS in March 2019 for the THA/TKA complications measure, and the measure was revised accordingly. Available at: https://www.qualitynet.org/inpatient/measures/complication/methodology

LEJR = lower extremity joint replacement, LTCH = long-term care hospital, MS-DRG = Medicare Severity-Diagnosis Related Group, 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.

Exhibit F-2: Claims-based outcome definitions for the fracture analyses[6]

Domain	Outcome name	Definition	Measurement period(s)	Eligible sample
Medicare payments	Total Medicare standardized allowed amounts for the episode per days alive	The sum of Medicare payment and beneficiary out-of-pocket amounts for related items and services covered by Medicare Part A and Part B[7] during the measurement period that are included in the episode, divided by the number of days alive during the measurement period.	Anchor through days 0-89 PDP (episode), days 90-179 PDP (PEP #1), days 180-269 PDP (PEP #2), days 270-359 PDP (PEP #3)	Beneficiaries who: 1) receive an LEJR due to hip fracture; 2) have a complete FFS enrollment history six months prior to the anchor hospitalization; 3) have consistent, reliable sex and age data (age <115); 4) maintain Parts A and B enrollment throughout the measurement period; 5) have a 90-day episode period that ends on or before December 31, 2019; 6) have non-missing Medicare standardized allowed payment information for the episode.
	Medicare Part A standardized allowed amounts for the episode per days alive, by service	The sum of Medicare payment and beneficiary out-of-pocket amounts for readmissions, IRF, SNF, and hospice services covered under Medicare Part A during the measurement period, divided by the number of days alive during the measurement period.	Days 0-89 PDP (episode), days 90-179 PDP (PEP #1), days 180-269 PDP (PEP #2), days 270-359 PDP (PEP #3)	Beneficiaries who: 1) receive an LEJR due to hip fracture; 2) have a complete FFS enrollment history six months prior to the anchor hospitalization; 3) have consistent, reliable sex and age data (age <115); 4) maintain Parts A and B enrollment throughout the measurement period; 5) have a 90-day episode period that ends on or before December 31, 2019; 6) have non-missing Medicare standardized allowed payment information for the episode.

---

[6]   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. For the fracture analyses, we ran outcomes separately for three different samples: 1) meet all eligibility criteria and exclude deaths during the episode (per the Final Rule); 2) meet all eligibility criteria and include deaths during the episode; and 3) meet all eligibility criteria and exclude deaths that occurred during the post-episode period (e.g., exclude episodes from the 90-179 PDP analysis if the beneficiary passed away during that period).

[7]   Episode-related items and services paid under Medicare Part A or Part B, after exclusions are applied, include: physician services; inpatient hospital services (including readmissions with certain exceptions discussed in the Final Rule); IPF services; LTCH 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, cont'd	Medicare standardized allowed amounts for HHA services for the episode per days alive	The sum of Medicare payment and beneficiary out-of-pocket amounts for HH services covered under Medicare Parts A and B during the measurement period, divided by the number of days alive during the measurement period.	Days 0-89 PDP (episode), days 90-179 PDP (PEP #1), days 180-269 PDP (PEP #2), days 270-359 PDP (PEP #3)	Beneficiaries who: 1) receive an LEJR due to hip fracture; 2) have a complete FFS enrollment history six months prior to the anchor hospitalization; 3) have consistent, reliable sex and age data (age <115); 4) maintain Parts A and B enrollment throughout the measurement period; 5) have a 90-day episode period that ends on or before December 31, 2019; 6) have non-missing Medicare standardized allowed payment information for the episode.
	Medicare Part B standardized allowed amounts for DME for the episode per days alive	The sum of Medicare payment and beneficiary out-of-pocket amounts for DME covered under Medicare Part B during the measurement period, divided by the number of days alive during the measurement period.	Days 0-89 PDP (episode), days 90-179 PDP (PEP #1), days 180-269 PDP (PEP #2), days 270-359 PDP (PEP #3)	Beneficiaries who: 1) receive an LEJR due to hip fracture; 2) have a complete FFS enrollment history six months prior to the anchor hospitalization; 3) have consistent, reliable sex and age data (age <115); 4) maintain Parts A and B enrollment throughout the measurement period; 5) have a 90-day episode period that ends on or before December 31, 2019; 6) have non-missing Medicare standardized allowed payment information for the episode.
Utilization	Any IRF	The proportion of episodes with beneficiaries using any IRF services during the measurement period, as indicated by non-zero Medicare payment and beneficiary out-of-pocket amounts for IRF services covered under Medicare Part A.	Days 0-89 PDP (episode), days 90-179 PDP (PEP #1), days 180-269 PDP (PEP #2), days 270-359 PDP (PEP #3)	Beneficiaries who: 1) receive an LEJR due to hip fracture; 2) have a complete FFS enrollment history six months prior to the anchor hospitalization; 3) have consistent, reliable sex and age data (age <115); 4) maintain Parts A and B enrollment throughout the measurement period; 5) have a 90-day episode period that ends on or before December 31, 2019; 6) have non-missing Medicare standardized allowed payment information for the episode.
	Any SNF	The proportion of episodes with beneficiaries using any SNF services during the measurement period, as indicated by non-zero Medicare payment and beneficiary out-of-pocket amounts for SNF services covered under Medicare Part A.	Days 0-89 PDP (episode), days 90-179 PDP (PEP #1), days 180-269 PDP (PEP #2), days 270-359 PDP (PEP #3)	Beneficiaries who: 1) receive an LEJR due to hip fracture; 2) have a complete FFS enrollment history six months prior to the anchor hospitalization; 3) have consistent, reliable sex and age data (age <115); 4) maintain Parts A and B enrollment throughout the measurement period; 5) have a 90-day episode period that ends on or before December 31, 2019; 6) have non-missing Medicare standardized allowed payment information for the episode.

Domain	Outcome name	Definition	Measurement period(s)	Eligible sample
Utilization, cont'd	Any NF	The proportion of episodes with beneficiaries using any NF services during the measurement period, as indicated by a completed assessment in the MDS, or using any SNF services during the measurement period, as indicated by non-zero Medicare payment and beneficiary out-of-pocket amounts for SNF services covered under Medicare Part A.	Days 0-89 PDP (episode), days 90-179 PDP (PEP #1), days 180-269 PDP (PEP #2), days 270-359 PDP (PEP #3)	Beneficiaries who: 1) receive an LEJR due to hip fracture; 2) have a complete FFS enrollment history six months prior to the anchor hospitalization; 3) have consistent, reliable sex and age data (age <115); 4) maintain Parts A and B enrollment throughout the measurement period; 5) have a 90-day episode period that ends on or before December 31, 2019; 6) have non-missing Medicare standardized allowed payment information for the episode.
	Any HH	The proportion of episodes with beneficiaries using any HHA services during the measurement 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.	Days 0-89 PDP (episode), days 90-179 PDP (PEP #1), days 180-269 PDP (PEP #2), days 270-359 PDP (PEP #3)	Beneficiaries who: 1) receive an LEJR due to hip fracture; 2) have a complete FFS enrollment history six months prior to the anchor hospitalization; 3) have consistent, reliable sex and age data (age <115); 4) maintain Parts A and B enrollment throughout the measurement period; 5) have a 90-day episode period that ends on or before December 31, 2019; 6) have non-missing Medicare standardized allowed payment information for the episode.
	Any hospice	The proportion of episodes with beneficiaries using any hospice services during the measurement period, as indicated by non-zero Medicare payment and beneficiary out-of-pocket amounts for hospice services covered under Medicare Part A.	Days 0-89 PDP (episode), days 90-179 PDP (PEP #1), days 180-269 PDP (PEP #2), days 270-359 PDP (PEP #3)	Beneficiaries who: 1) receive an LEJR due to hip fracture; 2) have a complete FFS enrollment history six months prior to the anchor hospitalization; 3) have consistent, reliable sex and age data (age <115); 4) maintain Parts A and B enrollment throughout the measurement period; 5) have a 90-day episode period that ends on or before December 31, 2019; 6) have non-missing Medicare standardized allowed payment information for the episode.
	Any PT/OT	The proportion of episodes with beneficiaries using any outpatient PT/OT services during the measurement period, as indicated by having ≥1 outpatient PT/OT visits covered under Medicare Part B during the 90-day post-discharge period.	Days 0-89 PDP (episode), days 90-179 PDP (PEP #1), days 180-269 PDP (PEP #2), days 270-359 PDP (PEP #3)	Beneficiaries who: 1) receive an LEJR due to hip fracture; 2) have a complete FFS enrollment history six months prior to the anchor hospitalization; 3) have consistent, reliable sex and age data (age <115); 4) maintain Parts A and B enrollment throughout the measurement period; 5) have a 90-day episode period that ends on or before December 31, 2019; 6) have non-missing Medicare standardized allowed payment information for the episode.

Domain	Outcome name	Definition	Measurement period(s)	Eligible sample
Utilization, cont'd	Days in community	Count of days during the measurement period that the beneficiary did not receive health care services (IPPS, SNF, IRF, LTCH, hospice, other inpatient, or HH) covered under Medicare Part A or B and did not use any NF services, as indicated by a completed assessment in MDS, during the measurement period.	Days 0-89 PDP (episode), days 90-179 PDP (PEP #1), days 180-269 PDP (PEP #2), days 270-359 PDP (PEP #3)	Beneficiaries who: 1) receive an LEJR due to hip fracture; 2) have a complete FFS enrollment history six months prior to the anchor hospitalization; 3) have consistent, reliable sex and age data (age <115); 4) maintain Parts A and B enrollment throughout the measurement period; 5) have a 90-day episode period that ends on or before December 31, 2019; 6) have non-missing Medicare standardized allowed payment information for the episode.
DME	Use of bathroom aids	The proportion of episodes with beneficiaries using any bathroom aids covered under by Medicare Part B during the measurement period.	Days 0-89 PDP (episode), days 0-359 (one year)	Beneficiaries who: 1) receive an LEJR due to hip fracture; 2) have a complete FFS enrollment history six months prior to the anchor hospitalization; 3) have consistent, reliable sex and age data (age <115); 4) maintain Parts A and B enrollment throughout the measurement period; 5) have a 90-day episode period that ends on or before December 31, 2019; 6) have non-missing Medicare standardized allowed payment information for the episode.
	Use of mobility aids (canes and walkers)	The proportion of episodes with beneficiaries using any mobility aids (canes, walkers) covered under by Medicare Part B during the measurement period.	Days 0-89 PDP (episode), days 0-359 (one year)	Beneficiaries who: 1) receive an LEJR due to hip fracture; 2) have a complete FFS enrollment history six months prior to the anchor hospitalization; 3) have consistent, reliable sex and age data (age <115); 4) maintain Parts A and B enrollment throughout the measurement period; 5) have a 90-day episode period that ends on or before December 31, 2019; 6) have non-missing Medicare standardized allowed payment information for the episode.
	Use of wheelchairs	The proportion of episodes with beneficiaries using any wheelchairs covered under by Medicare Part B (either rented or purchased) during the measurement period.	Days 0-89 PDP (episode), days 90-179 PDP (PEP #1), days 180-269 PDP (PEP #2), days 270-359 PDP (PEP #3)	Beneficiaries who: 1) receive an LEJR due to hip fracture; 2) have a complete FFS enrollment history six months prior to the anchor hospitalization; 3) have consistent, reliable sex and age data (age <115); 4) maintain Parts A and B enrollment throughout the measurement period; 5) have a 90-day episode period that ends on or before December 31, 2019; 6) have non-missing Medicare standardized allowed payment information for the episode.

Domain	Outcome name	Definition	Measurement period(s)	Eligible sample
DME, cont'd	Use of hospital beds	The proportion of episodes with beneficiaries using any hospital beds covered under by Medicare Part B (either rented or purchased) during the measurement period.	Days 0-89 PDP (episode), days 90-179 PDP (PEP #1), days 180-269 PDP (PEP #2), days 270-359 PDP (PEP #3)	Beneficiaries who: 1) receive an LEJR due to hip fracture; 2) have a complete FFS enrollment history six months prior to the anchor hospitalization; 3) have consistent, reliable sex and age data (age <115); 4) maintain Parts A and B enrollment throughout the measurement period; 5) have a 90-day episode period that ends on or before December 31, 2019; 6) have non-missing Medicare standardized allowed payment information for the episode.
	Use of patient lifts or transfer boards	The proportion of episodes with beneficiaries using any patient lifts or transfer boards covered under by Medicare Part B during the measurement period.	Days 0-89 PDP (episode), days 0-359 (one year)	Beneficiaries who: 1) receive an LEJR due to hip fracture; 2) have a complete FFS enrollment history six months prior to the anchor hospitalization; 3) have consistent, reliable sex and age data (age <115); 4) maintain Parts A and B enrollment throughout the measurement period; 5) have a 90-day episode period that ends on or before December 31, 2019; 6) have non-missing Medicare standardized allowed payment information for the episode.
Quality	Unplanned readmission rate	The proportion of episodes with one or more unplanned readmissions for any eligible condition during the measurement period. This measure was based on specifications for the NQF-endorsed all-cause unplanned readmission measure (NQF measure 1789). Following these specifications, we excluded planned admissions, based on AHRQ Clinical Classification System Procedure and Diagnoses codes.	Days 0-89 PDP (episode), days 90-179 PDP (PEP #1), days 180-269 PDP (PEP #2), days 270-359 PDP (PEP #3)	Beneficiaries who: 1) receive an LEJR due to hip fracture; 2) have a complete FFS enrollment history six months prior to the anchor hospitalization; 3) have consistent, reliable sex and age data (age <115); 4) maintain Parts A and B enrollment throughout the measurement period; 5) have a 90-day episode period that ends on or before December 31, 2019; 6) have non-missing Medicare standardized allowed payment information for the episode; 7) are discharged from the anchor hospitalization in accordance with medical advice.
	Emergency department visit rate	The proportion of episodes with one or more ED visits during the measurement 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 emergency department but was not admitted to the hospital.	Days 0-89 PDP (episode), days 90-179 PDP (PEP #1), days 180-269 PDP (PEP #2), days 270-359 PDP (PEP #3)	Beneficiaries who: 1) receive an LEJR due to hip fracture; 2) have a complete FFS enrollment history six months prior to the anchor hospitalization; 3) have consistent, reliable sex and age data (age <115); 4) maintain Parts A and B enrollment throughout the measurement period; 5) have a 90-day episode period that ends on or before December 31, 2019; 6) have non-missing Medicare standardized allowed payment information for the episode; 7) are discharged from the anchor hospitalization in accordance with medical advice.

Domain	Outcome name	Definition	Measurement period(s)	Eligible sample
Quality, cont'd	All-cause mortality rate	Death from any cause during the measurement period.	Anchor through days 0-89 PDP (episode), days 90-179 PDP (PEP #1), days 180-269 PDP (PEP #2), days 270-359 PDP (PEP #3)	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 measurement period. Beneficiaries who: 1) receive an LEJR due to hip fracture; 2) have a complete FFS enrollment history six months prior to the anchor hospitalization; 3) have consistent, reliable sex and age data (age <115); 4) maintain Parts A and B enrollment throughout the measurement period; 5) have a 90-day episode period that ends on or before December 31, 2019; 6) have non-missing Medicare standardized allowed payment information for the episode; 7) are discharged from the anchor hospitalization in accordance with medical advice; 8) have not received hospice care in the six months prior to admission.
	Incidence of any complications[8]	The proportion of episodes with incidence (during the anchor hospitalization or a readmission) of: AMI, pneumonia, or sepsis/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.	Anchor through days 0-89 PDP (episode)	Beneficiaries who: 1) receive an LEJR due to hip fracture; 2) have a complete FFS enrollment history six months prior to the anchor hospitalization; 3) have consistent, reliable sex and age data (age <115); 4) maintain Parts A and B enrollment throughout the measurement period; 5) have a 90-day episode period that ends on or before December 31, 2019; 6) have non-missing Medicare standardized allowed payment information for the episode; 7) are discharged from the anchor hospitalization in accordance with medical advice.

---

[8]   This measure was based on specifications for the NQF-endorsed THA/TKA complications measure among elective TKA/THA (NQF measure 1550). Death in the 30 days after discharge is part of the technical definition but is not included in our analysis.

LEWINGROUP®

Domain	Outcome name	Definition	Measurement period(s)	Eligible sample
Quality, cont'd	Incidence of severe pressure ulcers[9]	The proportion of fracture episodes with incidence of severe pressure ulcers (stage III, IV, or unstageable) during the measurement period. During the anchor hospitalization, the episode had a secondary ICD diagnosis code for pressure ulcers on any Medicare Part A claim, the anchor length of stay was three or more days, and pressure ulcers were not present on admission. After discharge, the episode had a primary or secondary diagnosis of pressure ulcers on any Medicare Part A or Part B claim.	Anchor through days 0-89 PDP (episode), anchor through days 0-359 (one year)	Beneficiaries who: 1) receive an LEJR due to hip fracture; 2) have a complete FFS enrollment history six months prior to the anchor hospitalization; 3) have consistent, reliable sex and age data (age <115); 4) maintain Parts A and B enrollment throughout the measurement period; 5) have a 90-day episode period that ends on or before December 31, 2019; 6) have non-missing Medicare standardized allowed payment information for the episode.
	Incidence of delirium	The proportion of fracture episodes with incidence of delirium during the measurement period. During the anchor hospitalization, the episode had a secondary ICD diagnosis code for delirium on any Medicare Part A claim and delirium was not present on admission. After discharge, the episode had a primary or secondary diagnosis of delirium on any Medicare Part A or Part B claim.	Anchor through days 0-89 PDP (episode), anchor through days 0-359 (one year)	Beneficiaries who: 1) receive an LEJR due to hip fracture; 2) have a complete FFS enrollment history six months prior to the anchor hospitalization; 3) have consistent, reliable sex and age data (age <115); 4) maintain Parts A and B enrollment throughout the measurement period; 5) have a 90-day episode period that ends on or before December 31, 2019; 6) have non-missing Medicare standardized allowed payment information for the episode.

Note: AMI = acute myocardial infarction, HH = home health, IPPS = inpatient prospective payment system, IRF = inpatient rehabilitation facility, LTCH = long-term care hospital, MDS = Minimum Data Set, NF = nursing facility, OT = occupational therapy, PDP = post-discharge period, PEP = post-episode period, PT = physical therapy, SNF = skilled nursing facility, THA = total hip arthroplasty, TKA = total knee arthroplasty.

---

[9]   The specifications for this measure follow the pressure ulcer definition formerly used in the CMS Patient Safety Indicator (PSI) 90 measure.

LEWINGROUP®

Appendix G: Definitions of Patient Characteristics

Exhibit G-1: Patient characteristic variable definitions

Variable	Definition	Source
Age	Percent of patients by age category; 20 to 64, 65 to 79, 80 and above.	January 2012 - December 2014 (baseline) and April 2016 – September 2021 (intervention) Medicare Enrollment Database
Cancer	Percent of patients with cancer (HCC flag #8 – lymphoma and other cancers, #9 – colorectal, bladder, and other cancers, #10 – breast, prostate, and other cancers and tumors, #11 – metastatic cancer and acute leukemia, and #12 – lung and other severe cancers).	July 2011 – December 2014 (baseline) and April 2015 – September 2021 (intervention) Medicare Claims
Congestive heart failure	Percent of patients with congestive heart failure (HCC flag #85).	July 2011 – December 2014 (baseline) and April 2015 – September 2021 (intervention) Medicare Claims
Dementia	Percent of patients with dementia (with and without complications; HCC flags #51 and #52).	July 2011 – December 2014 (baseline) and April 2015 – September 2021 (intervention) Medicare Claims
Diabetes	Percent of patients with diabetes based on the CCW Chronic Conditions algorithm.[1]	July 2010 – December 2014 (baseline) and April 2014 – September 2021 (intervention) Medicare Claims
Disability, not due to ESRD	Percent disabled, based on Medicare eligibility status (not including ESRD).	January 2012 - December 2014 (baseline) and April 2016 – September 2021 (intervention) Medicare Enrollment Database
Eligible for Medicaid	Percent eligible for Medicaid based on Medicare enrollment file.	January 2012 - December 2014 (baseline) and April 2016 – September 2021 (intervention) Medicare Enrollment Database
Elective episode	Percent of patients *without* hip fractures at the anchor hospitalization based on ICD codes provided by CMMI on the CJR model website (https://innovation.cms.gov/Files/worksheets/cjr-icd10hipfracturecodes.xlsx).	January 2012 - December 2014 (baseline) and April 2016 – September 2021 (intervention) Medicare Claims
Fracture status	Percent of patients *with* hip fractures at the anchor hospitalization based on ICD codes provided by CMMI on the CJR model website (https://innovation.cms.gov/Files/worksheets/cjr-icd10hipfracturecodes.xlsx).	January 2012 - December 2014 (baseline) and April 2016 – September 2021 (intervention) Medicare Claims

---

[1] Chronic Conditions Data Warehouse. Other Chronic Conditions. https://www2.ccwdata.org/web/guest/condition-categories-chronic

Variable	Definition	Source
HCC score	Average CMS-HCC score that corresponds to the HCCs present during the one year prior to the anchor hospitalization. HCC scores of less than 1.0 indicate the patient is healthier than the average Medicare beneficiary, while scores greater than 1.0 indicate a patient is unhealthier than the average Medicare beneficiary.	January 2012 - December 2014 (baseline) and April 2016 – September 2021 (intervention) Medicare Claims
Hypertension	Percent of patients with hypertension based on the CCW Chronic Conditions algorithm.[1]	July 2011 – December 2014 (baseline) and April 2015 – September 2021 (intervention) Medicare Claims
Length of stay	Mean length of anchor hospitalization in days. Number of days between the date of the procedure and the date of discharge from the hospital.	July 2011 – December 2014 (baseline) and April 2015 – September 2021 (intervention) Medicare Claims
MS-DRG 469	Percent of patients discharged under MS-DRG 469 (major joint replacement or reattachment of lower extremity with major complications or comorbidities) for the anchor hospitalization.	January 2012 - December 2014 (baseline) and April 2016 – September 2021 (intervention) Medicare Claims
MS-DRG 470	Percent of patients discharged under MS-DRG 470 (major joint replacement or reattachment of lower extremity without major complications or comorbidities) for the anchor hospitalization.	January 2012 - December 2014 (baseline) and April 2016 – September 2021 (intervention) Medicare Claims
MS-DRG 521	Percent of patients discharged under MS-DRG 521 (hip replacement with principal diagnosis of hip fracture with major complications or comorbidities) for the anchor hospitalization. MS-DRG 521 took effect October 2020.	January 2012 - December 2014 (baseline) and April 2016 – September 2021 (intervention) Medicare Claims
MS-DRG 522	Percent of patients discharged under MS-DRG 522 (hip replacement with principal diagnosis of hip fracture without major complications or comorbidities) for the anchor hospitalization. MS-DRG 522 took effect October 2020.	January 2012 - December 2014 (baseline) and April 2016 – September 2021 (intervention) Medicare Claims
Obesity	Percent of patients obese or with a BMI of greater than 30 based on the CCW Chronic Conditions algorithm.[2]	July 2010 – December 2014 (baseline) and April 2014 – September 2021 (intervention) Medicare Claims
Prior care use	Percent of patients with any care use (inpatient, SNF, IRF, HH, or LTCH) during the six months prior to anchor hospitalization.	July 2011 - December 2014 (baseline) and October 2015 – September 2021 (intervention) Medicare Claims
Prior HH use	Percent of patients with one or more instances of HH use during the six months prior to the anchor hospitalization.	July 2011 - December 2014 (baseline) and October 2015 – September 2021 (intervention) Medicare Claims

---

[2] Chronic Conditions Data Warehouse. Other Chronic Health, Mental Health, and Potentially Disabling Conditions. https://www2.ccwdata.org/web/guest/condition-categories-other

Variable	Definition	Source
Prior inpatient ACH stay	Percent of patients with one or more inpatient acute care hospitalizations during the six months prior to the anchor hospitalization.	July 2011 - December 2014 (baseline) and October 2015 – September 2021 (intervention) Medicare Claims
Prior IRF use	Percent of patients with one or more IRF stays during the six months prior to the anchor hospitalization.	July 2011 - December 2014 (baseline) and October 2015 – September 2021 (intervention) Medicare Claims
Prior other inpatient use	Percent of patients with one or more stays in other inpatient settings (psychiatric hospital, cancer center) during the six months prior to the anchor hospitalization.	July 2011 - December 2014 (baseline) and October 2015 – September 2021 (intervention) Medicare Claims
Prior SNF use	Percent of patients with one or more SNF stays during the six months prior to the anchor hospitalization.	July 2011 - December 2014 (baseline) and October 2015 – September 2021 (intervention) Medicare Claims
Race or ethnicity	Percent of patients by race or ethnicity: Black or African American, Hispanic, other race or ethnicity, unknown race or ethnicity, or white.	January 2012 - December 2014 (baseline) and April 2016 – September 2021 (intervention) Medicare Enrollment Database
Sex	Percent of female patients.	January 2012 - December 2014 (baseline) and April 2016 – September 2021 (intervention) Medicare Enrollment Database
Tobacco	Percent of patients with any tobacco use based on the CCW Chronic Conditions algorithm.[2]	July 2011 - December 2014 (baseline) and October 2015 – September 2021 (intervention) Medicare Claims
Total knee arthroplasty	Percent of patients with a TKA based on ICD diagnosis codes indicating a knee replacement.	July 2011 - December 2014 (baseline) and October 2015 – September 2021 (intervention) Medicare Claims

Note: ACH = acute care hospital, BMI = body mass index, CCW = Chronic Condition Warehouse, CMMI = Center for Medicare & Medicaid Innovation, CMS = Centers for Medicare & Medicaid Services, ESRD = end-stage renal disease, HCC = hierarchical condition category, HH = home health, ICD = International Classification of Diseases, IRF = inpatient rehabilitation facility, LTCH = long-term care hospital, MS-DRG = Medicare Severity-Diagnosis Related Group, SNF = skilled nursing facility, TKA = total knee arthroplasty.

Appendix H: Care Coordination Survey Questions

Care Coordination Survey for the Comprehensive Care for Joint Replacement (CJR) Model

Please contact the CJR Evaluation Team (CJREval@lewin.com) for information about this survey.

Survey Instructions

As part of the Centers for Medicare & Medicaid Services' (CMS) evaluation of the Comprehensive Care for Joint Replacement (CJR) model, we are asking all CJR participant hospitals to answer a few quick questions about their care coordination experiences with the CJR model. By participating in the Care Coordination Survey, you have a unique opportunity to share your insights and experiences directly with CMS.

This survey is designed to help CMS understand the variety of care coordination activities and strategies that CJR participant hospitals implement for Medicare beneficiaries who receive lower extremity joint replacement (or LEJR) surgery. Please complete all questions in the CJR Care Coordination Survey and submit your responses online no later than **August 18, 2020.** We encourage you to collaborate with individuals who also conduct care coordination activities in your hospital on this survey, submitting a single response based on your collaborative discussion. **Note:** if you represent multiple hospitals, please respond only about the hospital for which the CJR hospital representative selected you. Your participation in this survey is greatly appreciated!

Please review the **following tips** for completing the CJR Care Coordination Survey:

- Complete the survey on a computer instead of a mobile phone or tablet for ease of use.

- You may use the printable PDF version of the survey, found in the email message we sent you with the survey link on July 28, 2020, to gather and document responses before starting the online survey tool if it is helpful; this is not a requirement.

- At the beginning of the survey, you will be asked to include your hospital's name, zip code, and CMS Certification Number (CCN). If you do not know your hospital's CCN, you may find it in the email message we sent you with the survey link on July 28, 2020.

- Click the "Save and Next" button to move to the next page. Click "Previous" to return to the previous page.

- If you are unable to complete the survey in one sitting, you may return to your survey responses once you have closed your browser window. You must click the "Save and Next" button at the bottom of the page to save your responses up to that point.

- Use the same computer and web browser to return to your saved survey responses.

- To complete the survey, you must click the "Submit" button on the last page of the survey to submit your results.

If you have any questions regarding the CJR Care Coordination Survey, please contact the CJR Evaluation team at CJReval@lewin.com.

Care Coordination Survey

Section 1: Hospital and Care Coordinator Information

1)   Please provide the name of your hospital.

_____

_____

2)   Please provide your hospital's zip code.

_____

_____

3)   Please provide your hospital's CMS Certificate Number (CCN).[1]

_____

_____

4)   Please select the **profession or occupation** that best describes you:
- o   Nurse
- o   Nurse Practitioner
- o   Occupational Therapist
- o   Physical Therapist
- o   Physician Assistant
- o   Social Worker
- o   Other (please describe) _____

_____

5)   Please select the **title** that best describes you:
- o   Care Coordinator
- o   Case Manager
- o   Health Educator
- o   Navigator
- o   Other (please describe) _____

_____

---

[1] Note: If you do not know your hospital's CCN, you may find it in the email message with the survey link that was sent to you on July 28, 2020.

6) How many years have you been working in **a care coordination role** at your hospital?
- o Less than one year
- o 1 – 4 years
- o 5 – 9 years
- o More than 10 years

7) Did your hospital **hire you or assign you** to perform care coordination activities as a result of the CJR model?
- o Yes, due to the CJR model
- o No, not due to the CJR model
- o Don't know

8) How much of **your time** is dedicated to care coordination activities?
- o 0 – 25%
- o 26 – 50%
- o 51 – 75%
- o 76 – 100%

9) **In addition to your role**, has your hospital dedicated staff or other resources to care coordination?
- o Yes, due to the CJR model
- o Yes, but not due to the CJR model
- o No
- o Don't know

10) If yes, your hospital has dedicated staff or other resources to care coordination in addition to your role, please describe (e.g., number of staff members allocated to care coordination, implementation of patient tracking tools, etc.).

Section 2: Care Coordination Activities

11) How **influential** was the CJR model in your hospital's decision to **implement or enhance** the following activities for hip and knee replacement surgery patients?

Activities	Very influential	Somewhat influential	Not at all influential	Don't know	Not doing this activity
Perform multidisciplinary rounds					
Develop discharge or transition plans (e.g., location and expected length of stay)					
Coordinate the patient's durable medical equipment (DME)					
Conduct medication reconciliation					
Track patient referrals during the 90-day post-hospital discharge period					
Track patient outcomes during the 90-day post-hospital discharge period					

12) How **influential** was the CJR model in your hospital's decision to **implement or enhance** the following activities for hip and knee replacement surgery patients?

Activities	Very influential	Somewhat influential	Not at all influential	Don't know	Not doing this activity
Hold a preoperative one-on-one meeting or phone call with the patient about patient status and discharge plan					
Provide pre-surgical education (e.g., "joint class" or "joint camp") to patients					
Actively engage the patient in developing the care and discharge plans					
Provide the patient with a primary point of contact on the healthcare team (e.g., nurse, navigator)					
Follow up with the patient post-discharge					

13) How **influential** was the CJR model in your hospital's decision to **implement or enhance** the following activities for hip and knee replacement surgery patients?

Activities	Very influential	Somewhat influential	Not at all influential	Don't know	Not doing this activity
Implement a formal process for creating a preferred provider network					
Hold regular meetings with post-acute care providers (e.g., to discuss referrals, hospital expectations)					
Communicate with post-acute care providers regarding individual patient care and patient status					
Share patient data as part of referral process with post-acute care providers					
Monitor patient status while patients receive post-acute care services					
Receive patient data from post-acute care providers					

14) How **influential** was the CJR model in your hospital's decision to **implement or enhance** the following activities for hip and knee replacement surgery patients?

Activities	Very influential	Somewhat influential	Not at all influential	Don't know	Not doing this activity
Perform risk assessment or stratification					
Collect and analyze patient reported outcomes (PRO) data					
Analyze Medicare claims data					
Review readmissions data					

15) Please describe any additional activities that your hospital has implemented or enhanced to support care for hip and knee replacement surgery patients as a result of the CJR model (which began April 1, 2016).

Section 3: Patient Selection and Transition of Care

16) Reflecting on your hospital's hip and knee replacement surgery patients who go to skilled nursing facilities or inpatient rehabilitation facilities, about how often were the following characteristics a **determining factor** in the discharge destination decision?

Patient Characteristics	Rarely (less than 10%)	Occasionally (about 30%)	Sometimes (about 50%)	Frequently (about 70%)	Usually (about 90% or more)
Age					
Comorbidities					
Mental health					
Physical condition					
Cognitive functioning					
Type of joint replacement surgery (e.g., hip, knee, elective, fracture)					
Access to accessible housing					
Availability of a caregiver					
Food insecurity/ access to nutritious foods					
Insurance type (e.g., private payer, Medicaid)					
Lack of quality post-acute care providers					
Language					
Transportation availability					

17) Please describe any additional factors considered in the discharge destination decision for those hip and knee replacement surgery patients who go to skilled nursing facilities or inpatient rehabilitation facilities.

18) Reflecting on your hospital's hip and knee replacement surgery patients who go home, about how often were the following issues **a challenge** to a safe discharge home?

Category	Issues	Rarely (less than 10%)	Occasionally (about 30%)	Sometimes (about 50%)	Frequently (about 70%)	Usually (about 90% or more)
Social determinants of health	Access to accessible housing					
	Availability of a caregiver					
	Food insecurity/ access to nutritious foods					
	Transportation availability					

19) Please describe any additional issues that pose a challenge to a safe discharge home for hip and knee replacement surgery patients who go home following hip or knee replacement surgery.

_____

_____

20) Does your hospital provide a resource list of community services to hip and knee replacement surgery patients when they are discharged?

   o   Yes, provided to all patients
   o   Yes, provided to patients when needed
   o   No, none of the time
   o   Don't know

21) If yes, please describe the types of community services (e.g., housing resources, mental health services, adult daycare, home-delivered meal programs, transportation) provided on your hospital's resource list.

_____

_____

Section 4: Impact of Care Coordination

22) How does your hospital **measure your progress** towards your care coordination strategy? Select all that apply.

- o Analysis of electronic health record (EHR) data to track patient outcomes
- o Analysis of Medicare claims data to track patient outcomes
- o Collection and analysis of patient reported outcomes (PRO) data
- o Collection and analysis of patient satisfaction surveys
- o Data sharing and discussions with post-acute care providers
- o Data sharing and discussions with healthcare teams
- o Monthly status reports
- o Other (please describe)

  _____

  _____

23) Has your hospital identified **specific goals** for their care coordination activities? Some examples of goals that hospitals may have are the following: Using data to identify ways to prevent readmissions for other patients, or 100% of patients receive joint class education.

- o Yes
- o No
- o Don't know

24) If yes, what are those goals?

- o Goal 1: _____

  _____

- o Goal 2: _____

  _____

- o Goal 3: _____

  _____

25) Which **outcomes** are monitored to measure the success of your hospital's care coordination strategy? Select all that apply.

- o   Complications or infections

- o   Discharge disposition or status

- o   Emergency department use

- o   Hospital length of stay

- o   Net payment reconciliation amount (NPRA)

- o   Patient reported outcomes (PRO)

- o   Patients' physical progress after post-acute care or physical therapy

- o   Patient satisfaction

- o   Post-acute care utilization

- o   Readmissions

- o   Other (please describe)

_____

_____

Thank you very much.

We greatly appreciate your participation in this survey. Your participation will enable CMS to better understand the variety of care coordination activities and strategies that CJR participant hospitals implement for patients who receive hip and knee replacement surgeries.

Appendix I: Care Coordination Survey Results

Exhibit I-1: Frequencies for survey questions about how influential the CJR model was on hospitals' decisions to implement or enhance care coordination activities for hip and knee replacement surgery patients, CJR mandatory and opt-in hospitals

Activity (N=199 respondents)	Not doing or don't know % (n)	Not influential % (n)	Somewhat influential % (n)	Very influential % (n)
Perform multi-disciplinary rounds	13.1% (26)	29.1% (58)	37.7% (75)	20.1% (40)
Develop discharge plans				
Coordinate durable medical equipment	6.5% (13)	33.7% (67)	32.2% (64)	27.6% (55)
Conduct medication reconciliation	9.5% (19)	44.2% (88)	30.7% (61)	15.6% (31)
Track patient referrals during discharge period	17.6% (35)	8.0% (16)	24.1% (48)	50.3% (100)
Track patient outcomes during discharge period	11.1% (22)	8.5% (17)	23.1% (46)	57.3% (114)
Hold preoperative 1-1 meeting with patient	15.6% (31)	17.6% (35)	22.6% (45)	44.2% (88)
Provide pre-surgical education	6.5% (13)	24.1% (48)	25.1% (50)	44.2% (88)
Engage patient in developing discharge plan				
Provide patient with primary point of contact	8.0% (16)	19.1% (38)	24.6% (49)	48.2% (96)
Follow up with the post-discharge plan	7.0% (14)	13.6% (27)	28.1% (56)	51.3% (102)
Implement formal process for creating preferred provider networks	19.1% (38)	16.1% (32)	26.6% (53)	38.2% (76)
Hold meetings with post-acute care providers	17.1% (34)	12.1% (24)	32.2% (64)	38.7% (77)
Communicate with post-acute care providers regarding patients	12.6% (25)	10.6% (21)	31.7% (63)	45.2% (90)
Share patient data with post-acute care providers	18.6% (37)	15.1% (30)	30.2% (60)	36.2% (72)
Monitor patient status while receiving post-acute care services	18.6% (37)	7.5% (15)	32.2% (64)	41.7% (83)
Receive patient data from post-acute care providers	25.1% (50)	11.6% (23)	25.1% (50)	38.2% (76)
Perform risk assessment or stratification	12.6% (25)	12.6% (25)	35.2% (70)	39.7% (79)
Collect and analyze patient reported data	17.1% (34)	7.5% (15)	23.6% (47)	51.8% (103)
Analyze Medicare claims data	18.6% (37)	7.5% (15)	26.1% (52)	47.7% (95)
Review readmissions data	8.0% (16)	14.6% (29)	30.2% (60)	47.2% (94)

Source: CJR evaluation team analysis of the care coordination survey (fielded July 28 through August 30, 2020). Shaded rows are suppressed due to small cell size (<10).

Exhibit I-2: **Frequencies for survey questions about how often the following characteristics were determining factors in the discharge destination decision to skilled nursing facilities or inpatient rehabilitation facilities for hip and knee replacement surgery patients, CJR mandatory and opt-in hospitals.**

Characteristic (N=199 respondents)	About 30% of the time or less % (n)	About 50% of the time % (n)	About 70% of the time or more % (n)
Age	58.3% (116)	19.1% (38)	22.6% (45)
Comorbidities	23.6% (47)	24.6% (49)	51.8% (103)
Mental health	64.3% (128)	18.1% (36)	17.6% (35)
Physical condition	11.6% (23)	15.6% (31)	72.9% (145)
Cognitive functioning	38.7% (77)	23.1% (46)	38.2% (76)
Type of joint replacement surgery (e.g., hip, knee, elective, fracture)	48.2% (96)	22.6% (45)	29.1% (58)
Access to accessible housing	54.3% (108)	21.6% (43)	24.1% (48)
Availability of a caregiver	26.1% (52)	22.6% (45)	51.3% (102)
Food insecurity/ access to nutritious foods	83.4% (166)	7.5% (15)	9% (18)
Insurance type (e.g., private payer, Medicaid)	71.4% (142)	16.1% (32)	12.6% (25)
Lack of quality post-acute care providers	76.9% (153)	11.6% (23)	11.6% (23)
Language			
Transportation availability			

Source: CJR evaluation team analysis of the care coordination survey (fielded July 28 through August 30, 2020). Shaded rows are suppressed due to small cell size (<10).

Exhibit I-3: **Frequencies for survey questions about how often the following social determinants of health are challenges to safe discharge home for those patients who go home for hip and knee replacement surgery patients, CJR mandatory and opt-in hospitals.**

Social determinant of health (N=199 respondents)	About 30% of the time or less % (n)	About 50% of the time % (n)	About 70% of the time or more % (n)
Access to accessible housing			
Availability of a caregiver	53.8% (107)	21.1% (42)	25.1% (50)
Food insecurity/ access to nutritious foods			
Transportation availability	81.9% (163)	9.6% (19)	8.5% (17)

Source: CJR evaluation team analysis of the care coordination survey (fielded July 28 through August 30, 2020). Shaded rows are suppressed due to small cell size (<10).

Exhibit I-4:   Frequency for survey question about if hospitals provide a list of community services to hip and knee replacement surgery patients when they are discharged, CJR mandatory and opt-in hospitals.

List of community services provided (N=199 respondents)	% (n)
Yes, provided to all patients	26.1% (52)
Yes, provided to patients when needed	56.8% (113)
No, none of the time	7.5% (15)
Don't know	9.6% (19)

Source: CJR evaluation team analysis of the care coordination survey (fielded July 28 through August 30, 2020).

Exhibit I-5:   Frequencies for survey question about how the hospitals measure progress towards their care coordination strategies for hip and knee replacement surgery patients (select all that apply), CJR mandatory and opt-in hospitals.

Methods to measure progress in care coordination (N=199 respondents)	% (n)
Analysis of electronic health record (EHR) data to track patient outcomes	72.9% (145)
Analysis of Medicare claims data to track patient outcomes	72.4% (144)
Collection and analysis of patient reported outcomes (PRO) data	66.3% (132)
Collection and analysis of patient satisfaction surveys	83.9% (167)
Data sharing and discussions with post-acute care providers	81.9% (163)
Data sharing and discussions with healthcare teams	66.8% (133)
Monthly status reports	62.8% (125)

Source: CJR evaluation team analysis of the care coordination survey (fielded July 28 through August 30, 2020).

Exhibit I-6:   Frequencies for survey question about which outcomes were monitored to measure success of the hospitals' care coordination strategies for hip and knee replacement surgery patients (select all that apply), CJR mandatory and opt-in hospitals.

Outcomes (N=199 respondents)	% (n)
Complications or infections	86.9% (173)
Discharge disposition or status	92.0% (183)
Emergency department use	60.8% (121)
Hospital length of stay	90.5% (180)
Net payment reconciliation amount (NPRA)	47.2% (94)
Patient reported outcomes (PRO)	61.3% (122)
Patients' physical progress after post-acute care or physical therapy	49.3% (98)
Patient satisfaction	87.4% (174)
Post-acute care utilization	75.9% (151)
Readmissions	97.0% (193)

Source: CJR evaluation team analysis of the care coordination survey (fielded July 28 through August 30, 2020).

Appendix J: Changes in Patient Characteristics

Composite Measure

Exhibit J-1: Average episode payments for the Elective Without MCC episode group decreased due to changes in patient mix under the CJR model
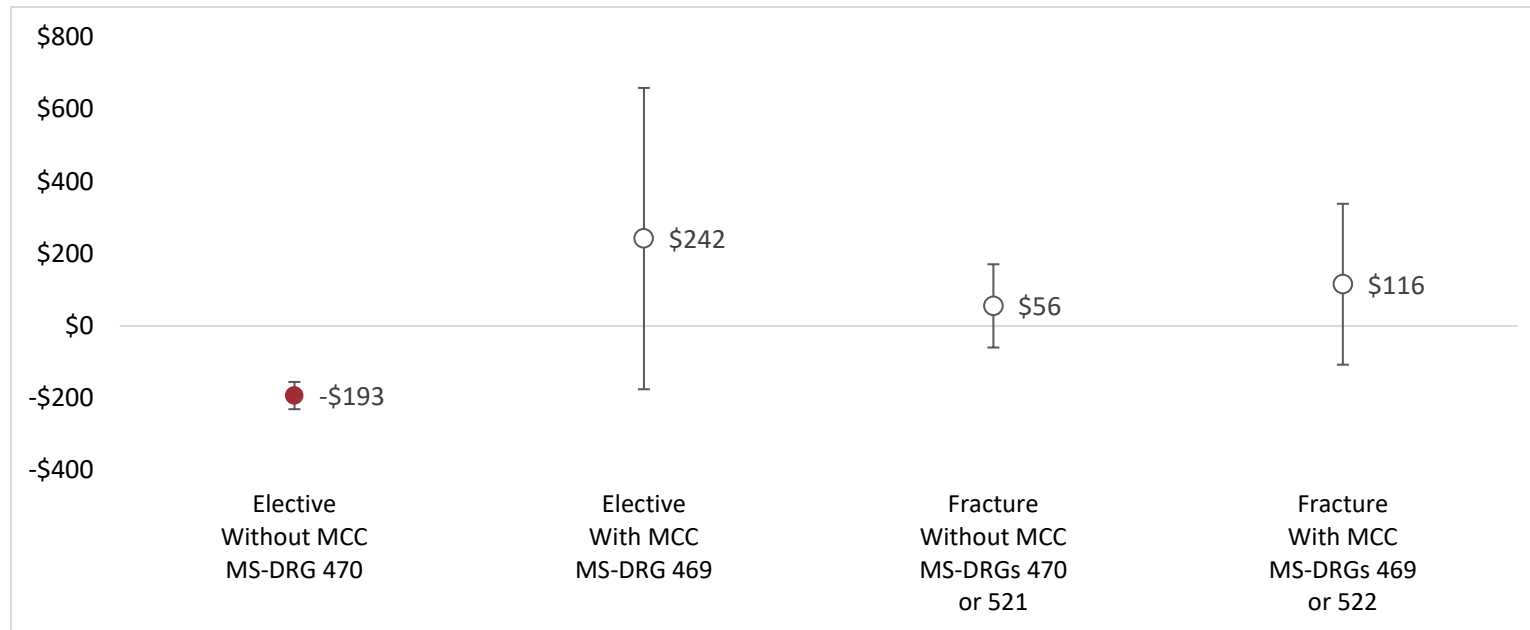
Source: CJR evaluation team analysis of Medicare claims and enrollment data for episodes initiated in 2012 through 2014 that ended between April 2012 and March 2015 (baseline) and episodes initiated during or after April 2016 that ended by September 2021 (intervention).

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

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

Elective episodes without MCC

Results by Performance Year

Composite Measure

Exhibit J-2: **For the Elective Without MCC episode group, the degree to which changes in the mix of patients decreased average episode payments grew over time**
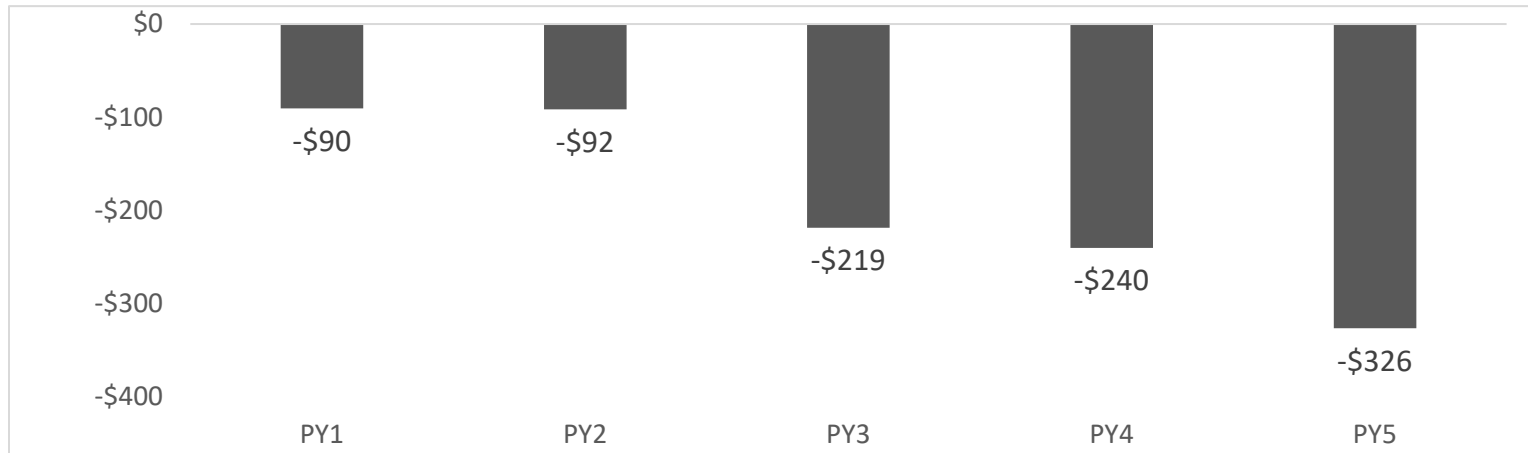
Source: CJR evaluation team analysis of Medicare claims and enrollment data for episodes initiated in 2012 through 2014 that ended between April 2012 and March 2015 (baseline) and episodes initiated during or after April 2016 that ended by September 2021 (intervention).

Notes: All estimates are obtained from a single analysis that measures how much of the relative change in total payments between CJR and control hospitals over each performance year (compared to the baseline) is attributable to relative changes in patient characteristics.

Individual Characteristics

Exhibit J-3: Change in patient characteristics by performance year, elective LEJR without MCC (MS-DRG 470)

Patient characteristics		Baseline Average		PY1		PY2		PY3		PY4		PY5	
		CJR	Control Group	DiD	P-value	DiD	P-value	DiD	P-value	DiD	P-value	DiD	P-value
Age	20-64	8.6%	8.9%	0.0	0.94	-0.1	0.82	-0.2	0.60	0.0	0.95	0.4	0.54
	65-79	70.1%	71.7%	-0.1	0.80	0.7	0.28	0.9	0.19	0.9	0.24	1.3	0.26
	80+	21.3%	19.4%	0.2	0.80	-0.5	0.36	-0.6	0.34	-0.9	0.13	-1.8	0.07
Sex	Female	64.6%	64.5%	-0.2	0.83	0.9	0.05	-0.6	0.22	-0.7	0.07	-0.2	0.71
Race and ethnicity	White	84.2%	87.7%	0.7	0.38	1.1	0.12	1.3	0.06	1.7	0.05	1.1	0.21
	Black or African American	7.0%	7.8%	-0.5	0.38	-0.7	0.16	-0.7	0.15	-0.2	0.63	0.2	0.74
	Hispanic	6.1%	3.1%	-0.3	0.63	-0.4	0.33	-0.7	0.16	-1.4	0.05	-1.3	0.11
	Other[a]	2.7%	1.5%	0.0	0.85	0.0	0.93	0.1	0.53	0.0	0.93	-0.1	0.83
Medicaid	Eligible for Medicaid	12.9%	10.2%	-0.9	0.25	-1.2	0.11	-1.4	0.04	-1.4	0.05	-1.0	0.27
Disability	Disability, no ESRD	16.6%	16.9%	-0.1	0.93	-0.3	0.68	-0.5	0.49	0.1	0.83	0.4	0.66
Health Status	HCC score	1.26	1.19	0.0	0.95	0.0	0.91	0.0	0.86	0.0	0.76	0.0	0.76
	Obesity	17.6%	18.1%	-0.9	0.61	0.4	0.86	1.0	0.69	0.5	0.84	0.9	0.73
	Diabetes	29.5%	27.3%	1.0	0.10	0.4	0.41	0.5	0.32	0.8	0.23	1.0	0.20
	Hypertension	75.2%	75.3%	-0.7	0.36	-0.4	0.51	0.3	0.76	-0.4	0.65	-0.4	0.65
	Dementia	3.2%	3.2%	0.2	0.23	-0.1	0.48	0.0	0.86	-0.1	0.57	-0.2	0.50
	CHF	12.4%	11.7%	-0.5	0.41	0.0	0.73	-1.0	0.15	-0.1	0.22	-0.2	0.17
Utilization in the six months prior to the anchor hospitalization	ACH stay	11.2%	11.2%	-0.6	0.07	-0.7	0.05	0.2	0.45	-0.1	0.64	0.2	0.65
	HH use	10.6%	9.8%	-0.7	0.19	-0.9	0.09	-0.2	0.71	-0.8	0.14	-0.4	0.52
	IRF stay	1.1%	1.1%	-0.3	0.11	-0.2	0.18	-0.1	0.66	-0.2	0.13	-0.2	0.40
	SNF stay	3.6%	3.2%	-0.1	0.80	-0.4	0.01	-0.5	<0.01	-0.4	<0.01	-0.4	0.06
	Any prior care	26.4%	26.2%	-1.0	0.14	-1.7	<0.01	-0.2	0.76	-0.8	0.10	-0.6	0.37

Source: CJR evaluation team analysis of Medicare claims and enrollment data for episodes initiated in 2012 through 2014 that ended between April 2012 and March 2015 (baseline) and episodes initiated during or after April 2016 that ended by September 2021 (intervention).

LEWINGROUP®

Notes: The estimates in this exhibit are the result of calculating the DiD of the unadjusted baseline and performance year averages for the CJR and control groups (net differences). DiD estimates that are significant at the 1%, 5%, or 10% significance level are indicated by red, orange, or yellow shaded cells, respectively.

Fracture is defined based on ICD codes for hip fracture provided by the CMMI on the CJR model website: https://innovation.cms.gov/Files/worksheets/cjr-icd10hipfracturecodes.xlsx.

ACH = acute care hospital, CHF = congestive heart failure, CI = confidence interval, CMMI = Center for Medicare & Medicaid Innovation, ESRD = end-stage renal disease, HCC = hierarchical condition category, HH = home health, ICD = International Classification of Diseases, IRF = inpatient rehabilitation facility, MCC = major complications or comorbidities, MS-DRG = Medicare Severity Diagnosis Related Group, pp = percentage point, PY = performance year, SNF = skilled nursing facility.

[a]   Other includes beneficiaries identified as "Asian," "American Indian/Alaska Native," or "Other."

Cumulative Results

Exhibit J-4: Change in patient characteristics, elective LEJRs without MCC (MS-DRG 470), PY1-5

Patient characteristics		CJR		Control group		Net differences (pp)	Net differences % of baseline	p-value	90% CI
		Baseline Average (N=119,239)	Intervention Average (N=167,133)	Baseline Average (N=146,644)	Intervention Average (N=172,997)				
Age	20-64	8.6%	6.4%	8.9%	6.7%	-0.3	-0.4%	0.93	-0.7 to 0.6
	65-79	70.1%	74.1%	71.7%	74.9%	0.8	1.2%	0.23	-0.3 to 1.9
	80+	21.3%	19.5%	19.4%	18.3%	-0.8	-3.7%	0.17	-1.7 to 0.2
Sex	Female	64.6%	63.7%	64.5%	63.7%	-0.2	-0.3%	0.61	-0.8 to 0.4
Race and ethnicity	White[a]	84.2%	86.0%	87.7%	88.2%	1.3	1.5%	0.03	0.3 to 2.2
	Black or African American[a]	7.0%	6.1%	7.8%	7.3%	-0.4	-5.5%	0.28	-1.0 to 0.2
	Hispanic[a]	6.1%	5.1%	3.1%	2.9%	-0.9	-14.6%	0.08	-1.7 to 0.0
	Other[ab]	2.7%	2.8%	1.5%	1.6%	0.0	0.1%	0.99	-0.3 to 0.3
Medicaid	Eligible for Medicaid	12.9%	9.9%	10.2%	8.4%	-1.3	-9.7%	0.01	-2.1 to -0.4
Disability	Disability, no ESRD	16.6%	14.8%	16.9%	15.2%	-0.1	-0.5%	0.87	-1.0 to 0.8
Health Status	HCC score	1.26	1.34	1.19	1.29	0.0	-1.8%	0.25	-0.1 to 0.0
	Obesity	17.6%	33.5%	18.1%	33.4%	0.7	3.8%	0.76	-2.9 to 4.3
	Diabetes	29.5%	29.5%	27.3%	26.5%	0.7	2.4%	0.11	0.0 to 1.4
	Hypertension	75.2%	77.0%	75.3%	77.4%	-0.3	-0.4%	0.66	-1.4 to 0.8
	Dementia	3.2%	2.8%	3.2%	2.8%	0.0	-1.5%	0.74	-0.3 to 0.2
	CHF	12.4%	12.1%	11.7%	11.7%	-0.3	-2.2%	0.47	-0.9 to 0.4
Utilization in the six months prior to the anchor hospitalization	ACH stay	11.2%	10.4%	11.2%	10.5%	-0.2	-1.7%	0.41	-0.6 to 0.2
	HH use	10.6%	9.4%	9.8%	9.2%	-0.6	-5.6%	0.22	-1.4 to 0.2
	IRF stay	1.1%	0.7%	1.1%	0.9%	-0.2	-16.3%	0.18	-0.4 to 0.0
	SNF stay	3.6%	2.6%	3.2%	2.6%	-0.4	-11.0%	<0.01	-0.6 to -0.2
	Any prior care	26.4%	24.9%	26.2%	25.7%	-0.9	-3.4%	0.04	-1.6 to -0.2

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 initiated during or after April 2016 that ended by September 2021 (intervention).

LEWINGROUP®

Notes: The estimates in this exhibit are the result of calculating the DiD of the unadjusted baseline and intervention averages for the CJR and control groups (net differences). DiD estimates that are significant at the 1%, 5%, or 10% significance level are indicated by red, orange, or yellow shaded cells, respectively.

Fracture is defined based on ICD codes for hip fracture provided by the CMMI on the CJR model website: https://innovation.cms.gov/Files/worksheets/cjr-icd10hipfracturecodes.xlsx.

ACH = acute care hospital, CHF = congestive heart failure, CI = confidence interval, CMMI = Center for Medicare & Medicaid Innovation, ESRD = end-stage renal disease, HCC = hierarchical condition category, HH = home health, ICD = International Classification of Diseases, IRF = inpatient rehabilitation facility, MCC = major complications or comorbidities, MS-DRG = Medicare Severity Diagnosis Related Group, pp = percentage point, PY = performance year, SNF = skilled nursing facility.

[a] The number of episodes for these measures is lower because episodes were dropped to account for missing observations. These measures are based on 118,676 CJR baseline episodes, 163,987 CJR intervention episodes, 146,020 control group baseline episodes, and 170,247 control group intervention episodes.

[b] Other includes beneficiaries identified as "Asian," "American Indian/Alaska Native," or "Other."

Trends over time

Exhibit J-5: Changes in the frequency of select patient characteristics in the Elective Without MCC episode group for mandatory hospitals from 2012 to 2021
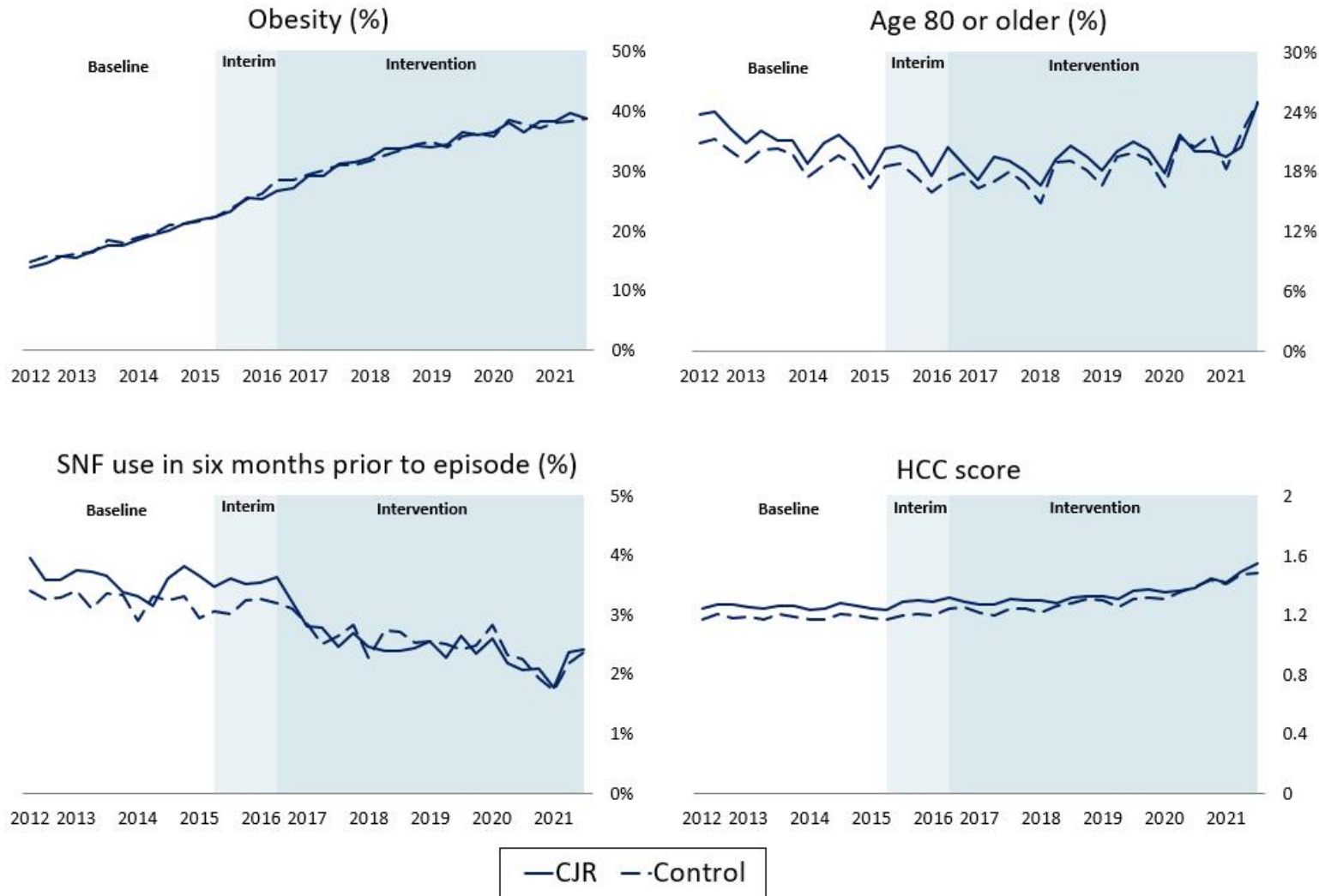
Source: CJR evaluation team analysis of Medicare claims and enrollment data for episodes that ended between April 2012 and September 2021 (intervention).

Elective episodes with MCC

Results by Performance Year

Exhibit J-6: Change in patient characteristics by performance year, elective LEJR with MCC (MS-DRG 469)

Patient characteristics		Baseline Average		PY1		PY2		PY3		PY4		PY5	
		CJR	Control Group	DiD	P-value	DiD	P-value	DiD	P-value	DiD	P-value	DiD	P-value
Age	20-64	12.0%	11.1%	-0.6	0.80	0.0	1.00	-0.8	0.65	-1.7	0.15	0.7	0.68
	65-79	54.4%	57.3%	-4.1	0.30	-4.7	0.10	-1.1	0.64	0.3	0.87	-2.9	0.21
	80+	33.7%	31.6%	4.6	0.27	4.7	0.07	1.9	0.35	1.3	0.49	2.3	0.29
Sex	Female	62.1%	61.2%	-1.4	0.67	1.0	0.75	1.8	0.48	0.0	0.99	-2.6	0.12
Race and ethnicity	White	82.7%	86.4%	0.1	0.96	-0.1	0.96	3.1	0.04	0.3	0.87	-0.8	0.55
	Black or African American	8.6%	8.9%	-3.0	0.16	-0.7	0.61	-2.3	0.07	-0.5	0.67	0.5	0.66
	Hispanic	6.2%	3.5%	2.1	0.04	0.6	0.69	0.4	0.76	1.0	0.36	0.5	0.50
	Other[a]	2.6%	1.2%	0.8	0.16	0.2	0.80	-1.2	0.04	-0.8	0.16	-0.2	0.65
Medicaid	Eligible for Medicaid	22.1%	16.4%	-2.6	0.33	-2.1	0.30	-6.0	<0.01	-3.6	0.02	-2.4	0.12
Disability	Disability, no ESRD	23.1%	21.3%	-1.1	0.69	0.8	0.71	-2.1	0.45	-3.5	0.07	-0.6	0.77
Health Status	HCC score	2.14	2.02	0.0	0.30	0.1	0.55	0.1	0.19	-0.1	0.76	0.0	0.77
	Obesity	21.7%	23.2%	-7.3	<0.01	4.1	0.15	1.6	0.56	1.3	0.70	-0.5	0.87
	Diabetes	38.8%	36.5%	-1.0	0.69	-2.4	0.25	-0.5	0.85	2.2	0.47	0.0	0.99
	Hypertension	82.2%	81.9%	-3.2	0.19	0.0	0.99	-3.0	0.19	-1.8	0.42	0.5	0.74
	Dementia	8.9%	9.3%	0.5	0.78	2.0	0.26	-1.5	0.29	-0.2	0.89	2.2	0.06
	CHF	27.5%	27.0%	-0.2	0.86	5.2	0.38	0.3	0.34	-0.7	0.37	2.3	0.83

Patient characteristics		Baseline Average		PY1		PY2		PY3		PY4		PY5	
		CJR	Control Group	DiD	P-value	DiD	P-value	DiD	P-value	DiD	P-value	DiD	P-value
Utilization in the six months prior to the anchor hospitalization	ACH stay	20.2%	18.5%	-3.7	0.13	1.8	0.43	0.9	0.68	-2.1	0.17	-0.2	0.94
	HH use	18.0%	18.1%	-2.4	0.50	4.0	0.13	0.0	0.99	1.7	0.29	0.8	0.74
	IRF stay	2.9%	2.4%	-1.7	0.05	-0.8	0.24	-0.2	0.82	0.2	0.83	-1.1	0.12
	SNF stay	8.7%	7.9%	0.2	0.94	0.1	0.95	2.0	0.23	-0.1	0.95	0.7	0.61
	Any prior care	40.6%	39.9%	-3.5	0.26	4.6	0.07	-0.3	0.92	-0.7	0.73	-2.6	0.28

Source: CJR evaluation team analysis of Medicare claims and enrollment data for episodes initiated in 2012 through 2014 that ended between April 2012 and March 2015 (baseline) and episodes initiated during or after April 2016 that ended by September 2021 (intervention).

Notes: The estimates in this exhibit are the result of calculating the DiD of the unadjusted baseline and performance year averages for the CJR and control groups (net differences). DiD estimates that are significant at the 1%, 5%, or 10% significance level are indicated by red, orange, or yellow shaded cells, respectively.

Fracture is defined based on ICD codes for hip fracture provided by the CMMI on the CJR model website: https://innovation.cms.gov/Files/worksheets/cjr-icd10hipfracturecodes.xlsx.

ACH = acute care hospital, CHF = congestive heart failure, CI = confidence interval, CMMI = Center for Medicare & Medicaid Innovation, ESRD = end-stage renal disease, HCC = hierarchical condition category, HH = home health, ICD = International Classification of Diseases, IRF = inpatient rehabilitation facility, MCC = major complications or comorbidities, MS-DRG = Medicare Severity Diagnosis Related Group, pp = percentage point, PY = performance year, SNF = skilled nursing facility.

a   Other includes beneficiaries identified as "Asian," "American Indian/Alaska Native," or "Other."

Cumulative Results

Exhibit J-7: Change in patient characteristics, elective LEJRs with MCC (MS-DRG 469), PY1-5

Patient characteristics		CJR		Control group		Net differences (pp)	Net differences % of baseline	p-value	90% CI
		Baseline Average (N=3,453)	Intervention Average (N=4,658)	Baseline Average (N=4,086)	Intervention Average (N=5,318)				
Age	20-64	12.0%	9.6%	11.1%	9.0%	-0.3	-2.8%	0.75	-2.1 to 1.4
	65-79	54.4%	59.1%	57.3%	64.4%	-2.4	-4.3%	0.15	-5.0 to 0.3
	80+	33.7%	31.3%	31.6%	26.6%	2.7	8.0%	0.11	-0.1 to 5.4
Sex	Female	62.1%	61.9%	61.2%	61.3%	-0.3	-0.6%	0.82	-2.8 to 2.2
Race and ethnicity	White[a]	82.7%	84.2%	86.4%	87.7%	0.3	0.4%	0.73	-1.2 to 1.9
	Black or African American[a]	8.6%	6.9%	8.9%	8.0%	-0.8	-9.2%	0.31	-2.0 to 0.5
	Hispanic[a]	6.2%	6.3%	3.5%	2.8%	0.8	13.4%	0.21	-0.3 to 1.9
	Other[ab]	2.6%	2.6%	1.2%	1.5%	-0.4	-14.5%	0.30	-1.0 to 0.2
Medicaid	Eligible for Medicaid	22.1%	15.6%	16.4%	13.1%	-3.2	-14.5%	0.00	-5.0 to -1.4
Disability	Disability, no ESRD	23.1%	21.7%	21.3%	21.2%	-1.3	-5.5%	0.38	-3.7 to 1.1
Health Status	HCC score	2.1	2.2	2.0	2.1	0.0	0.3%	0.90	-0.1 to 0.1
	Obesity	21.7%	36.4%	23.2%	37.4%	0.5	2.3%	0.84	-3.4 to 4.4
	Diabetes	38.8%	36.7%	36.5%	34.3%	0.1	0.2%	0.97	-3.1 to 3.2
	Hypertension	82.2%	83.9%	81.9%	84.6%	-1.0	-1.2%	0.49	-3.4 to 1.4
	Dementia	8.9%	7.9%	9.3%	7.6%	0.7	8.3%	0.46	-0.9 to 2.4
	CHF	27.5%	26.6%	27.0%	24.4%	1.6	5.9%	0.31	-1.0 to 4.3
Utilization in the six months prior to the anchor hospitalization	ACH stay	20.2%	19.7%	18.5%	18.3%	-0.3	-1.6%	0.82	-2.8 to 2.1
	HH use	18.0%	18.8%	18.1%	17.7%	1.2	6.7%	0.54	-2.0 to 4.4
	IRF stay	2.9%	2.6%	2.4%	2.7%	-0.6	-21.4%	0.17	-1.4 to 0.1
	SNF stay	8.7%	8.9%	7.9%	7.5%	0.6	7.1%	0.57	-1.2 to 2.4
	Any prior care	40.6%	40.0%	39.9%	39.7%	-0.4	-1.1%	0.81	-3.4 to 2.5

Source: CJR evaluation team analysis of Medicare claims and enrollment data for episodes initiated in 2012 through 2014 that ended between April 2012 and March 2015 (baseline) and episodes initiated during or after April 2016 that ended by September 2021 (intervention).

LEWINGROUP®

Notes: The estimates in this exhibit are the result of calculating the DiD of the unadjusted baseline and intervention averages for the CJR and control groups (net differences). DiD estimates that are significant at the 1% , 5% , or 10% significance level are indicated by red, orange, or yellow shaded cells, respectively.

The MS-DRG 470 is assigned at the anchor hospitalization discharge for major joint replacement or reattachment of lower extremity *without* MCC, while MS-DRG 469 is *with* MCC.

Fracture is defined based on ICD codes for hip fracture provided by the CMMI on the CJR model website: https://innovation.cms.gov/Files/worksheets/cjr-icd10hipfracturecodes.xlsx.

ACH = acute care hospital, CHF = congestive heart failure, CI = confidence interval, CMMI = Center for Medicare & Medicaid Innovation, ESRD = end-stage renal disease, HCC = hierarchical condition category, HH = home health, ICD = International Classification of Diseases, IRF = inpatient rehabilitation facility, MCC = major complications or comorbidities, MS-DRG = Medicare Severity Diagnosis Related Group, pp = percentage point, PY = performance year, SNF = skilled nursing facility.

[a]   The number of episodes for these measures is lower because episodes were dropped to account for missing observations. These measures are based on 3,441 CJR baseline episodes, 4,597 CJR intervention episodes, 4,075 control group baseline episodes, and 5,250 control group intervention episodes.

[b]   Other includes beneficiaries identified as "Asian," "American Indian/Alaska Native," or "Other."

Fracture episodes without MCC

Results by Performance Year

Exhibit J-8: Change in patient characteristics by performance year, fracture LEJRs without MCC (MS-DRGs 470 or 521)

Patient characteristics		Baseline Average		PY1		PY2		PY3		PY4		PY5	
		CJR	Control Group	DiD	P-value	DiD	P-value	DiD	P-value	DiD	P-value	DiD	P-value
Age	20-64	2.9%	3.2%	0.4	0.45	0.0	0.97	0.2	0.70	0.0	0.95	0.5	0.05
	65-79	29.5%	31.3%	-2.3	0.14	-2.0	0.01	0.4	0.78	0.1	0.89	-0.5	0.66
	80+	67.5%	65.5%	1.9	0.23	2.0	0.02	-0.6	0.64	-0.2	0.88	0.0	0.99
Sex	Female	75.2%	75.1%	-1.1	0.40	-1.5	0.13	-1.7	0.02	0.7	0.41	0.2	0.78
Race and ethnicity	White	88.4%	91.7%	-2.6	<0.01	0.2	0.87	0.6	0.52	0.4	0.69	0.0	0.95
	Black or African American	3.5%	4.9%	1.6	<0.01	0.4	0.45	-0.2	0.69	0.3	0.63	0.5	0.14
	Hispanic	5.1%	2.0%	0.5	0.46	-0.4	0.45	-0.6	0.40	-0.8	0.25	-0.9	0.05
	Other[a]	2.9%	1.3%	0.5	0.19	-0.2	0.65	0.2	0.43	0.1	0.80	0.4	0.20
Medicaid	Eligible for Medicaid	20.2%	16.4%	1.3	0.36	0.4	0.68	0.2	0.82	-0.7	0.35	-0.1	0.91
Disability	Disability, no ESRD	9.9%	10.6%	-1.1	0.26	-0.3	0.75	0.6	0.40	1.2	0.06	-0.1	0.87
Health Status	HCC score	2.32	2.27	0.0	0.45	0.1	0.73	0.0	0.42	0.0	0.54	0.0	0.09
	Obesity	4.1%	4.7%	-2.3	0.02	0.9	0.27	0.4	0.66	0.8	0.45	0.4	0.63
	Diabetes	28.9%	26.4%	2.2	0.12	1.5	0.17	-0.4	0.67	0.8	0.46	0.7	0.37
	Hypertension	74.6%	74.4%	-1.4	0.30	0.3	0.72	-0.9	0.45	-0.3	0.82	-0.8	0.44
	Dementia	29.2%	30.6%	1.0	0.43	2.8	0.01	1.0	0.41	0.0	0.99	2.4	<0.01
	CHF	23.6%	23.0%	2.0	0.25	0.8	0.04	-1.9	0.85	0.6	0.66	-0.2	0.89

Patient characteristics		Baseline Average		PY1		PY2		PY3		PY4		PY5	
		CJR	Control Group	DiD	P-value	DiD	P-value	DiD	P-value	DiD	P-value	DiD	P-value
Utilization in the six months prior to the anchor hospitalization	ACH stay	20.9%	21.3%	2.3	0.04	2.1	0.05	0.4	0.66	0.0	0.98	-0.4	0.58
	HH use	24.9%	24.9%	-0.7	0.60	0.8	0.52	-1.6	0.19	-1.2	0.40	-0.3	0.74
	IRF stay	2.5%	3.3%	0.7	0.15	0.9	0.09	-0.1	0.77	0.4	0.44	0.1	0.75
	SNF stay	11.5%	11.3%	-1.6	0.09	0.3	0.67	-0.7	0.25	-0.2	0.76	-0.5	0.29
	Any prior care	46.7%	48.2%	0.8	0.61	2.2	0.21	0.1	0.90	-0.1	0.96	0.9	0.34

Source: CJR evaluation team analysis of Medicare claims and enrollment data for episodes initiated in 2012 through 2014 that ended between April 2012 and March 2015 (baseline) and episodes initiated during or after April 2016 that ended by September 2021 (intervention).

Notes: The estimates in this exhibit are the result of calculating the DiD of the unadjusted baseline and performance year averages for the CJR and control groups (net differences). DiD estimates that are significant at the 1%, 5%, or 10% significance level are indicated by red, orange, or yellow shaded cells, respectively.

Fracture is defined based on ICD codes for hip fracture provided by the CMMI on the CJR model website: https://innovation.cms.gov/Files/worksheets/cjr-icd10hipfracturecodes.xlsx.

ACH = acute care hospital, CHF = congestive heart failure, CI = confidence interval, CMMI = Center for Medicare & Medicaid Innovation, ESRD = end-stage renal disease, HCC = hierarchical condition category, HH = home health, ICD = International Classification of Diseases, IRF = inpatient rehabilitation facility, MCC = major complications or comorbidities, MS-DRG = Medicare Severity Diagnosis Related Group, pp = percentage point, PY = performance year, SNF = skilled nursing facility.

[a] Other includes beneficiaries identified as "Asian," "American Indian/Alaska Native," or "Other."

Cumulative Results

Exhibit J-9: Change in patient characteristics, fracture LEJRs without MCC (MS-DRG 470 or 521), PY1-5

Patient characteristics		CJR		Control group		Net differences (pp)	Net differences % of baseline	p-value	90% CI
		Baseline Average (N=19,902)	Intervention Average (N=27,882)	Baseline Average (N=19,931)	Intervention Average (N=26,090)				
Age	20-64	2.9%	2.8%	3.2%	2.8%	0.3	8.7%	0.27	-0.1 to 0.6
	65-79	29.5%	33.6%	31.3%	36.0%	-0.6	-2.1%	0.43	-1.9 to 0.7
	80+	67.5%	63.6%	65.5%	61.2%	0.4	0.6%	0.65	-1.0 to 1.7
Sex	Female	75.2%	72.5%	75.1%	72.9%	-0.4	-0.6%	0.26	-1.1 to 0.2
Race and ethnicity	White[a]	88.4%	88.5%	91.7%	91.9%	0.0	0.0%	0.98	-1.0 to 1.0
	Black or African American[a]	3.5%	3.4%	4.9%	4.3%	0.4	11.6%	0.12	0.0 to 0.8
	Hispanic[a]	5.1%	4.9%	2.0%	2.4%	-0.6	-11.8%	0.17	-1.3 to 0.1
	Other[ab]	2.9%	3.2%	1.3%	1.4%	0.2	7.3%	0.48	-0.3 to 0.7
Medicaid	Eligible for Medicaid	20.2%	17.6%	16.4%	13.7%	0.0	0.2%	0.94	-0.9 to 1.0
Disability	Disability, no ESRD	9.9%	10.7%	10.6%	11.2%	0.2	2.0%	0.62	-0.5 to 0.9
Health Status	HCC score	2.3	2.4	2.3	2.4	0.0	0.4%	0.72	0.0 to 0.1
	Obesity	4.1%	8.9%	4.7%	9.1%	0.3	7.5%	0.67	-0.9 to 1.5
	Diabetes	28.9%	29.0%	26.4%	25.7%	0.8	2.8%	0.22	-0.3 to 1.9
	Hypertension	74.6%	76.9%	74.4%	77.2%	-0.6	-0.8%	0.53	-2.1 to 0.9
	Dementia	29.2%	26.8%	30.6%	26.7%	1.6	5.5%	0.00	0.7 to 2.5
	CHF	23.6%	22.1%	23.0%	21.6%	0.1	0.3%	0.93	-1.2 to 1.3
Utilization in the six months prior to the anchor hospitalization	ACH stay	20.9%	19.7%	21.3%	19.6%	0.5	2.4%	0.42	-0.5 to 1.5
	HH use	24.9%	22.7%	24.9%	23.2%	-0.6	-2.2%	0.57	-2.2 to 1.1
	IRF stay	2.5%	2.7%	3.3%	3.2%	0.3	12.5%	0.33	-0.2 to 0.9
	SNF stay	11.5%	10.1%	11.3%	10.4%	-0.4	-3.8%	0.25	-1.1 to 0.2
	Any prior care	46.7%	45.4%	48.2%	46.2%	0.8	1.6%	0.38	-0.7 to 2.2

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 initiated during or after April 2016 that ended by September 2021 (intervention).

LEWINGROUP®

Notes: The estimates in this exhibit are the result of calculating the DiD of the unadjusted baseline and intervention averages for the CJR and control groups (net differences). DiD estimates that are significant at the 1%, 5%, or 10% significance level are indicated by red, orange, or yellow shaded cells, respectively.

Fracture is defined based on ICD codes for hip fracture provided by the CMMI on the CJR model website: https://innovation.cms.gov/Files/worksheets/cjr-icd10hipfracturecodes.xlsx.

ACH = acute care hospital, CHF = congestive heart failure, CI = confidence interval, CMMI = Center for Medicare & Medicaid Innovation, ESRD = end-stage renal disease, HCC = hierarchical condition category, HH = home health, ICD = International Classification of Diseases, IRF = inpatient rehabilitation facility, MCC = major complications or comorbidities, MS-DRG = Medicare Severity Diagnosis Related Group, pp = percentage point, PY = performance year, SNF = skilled nursing facility.

[a]   The number of episodes for these measures is lower because episodes were dropped to account for missing observations. These measures are based on 19,865 CJR baseline episodes, 27,716 CJR intervention episodes, 19,903 control group baseline episodes, and 25,956 control group intervention episodes.

[b]   Other includes beneficiaries identified as "Asian," "American Indian/Alaska Native," or "Other."

Fracture episodes with MCC

Results by Performance Year

Exhibit J-10:   **Change in patient characteristics by performance year, fracture LEJRs with MCC (MS-DRGs 469 or 522)**

Patient characteristics		Baseline Average		PY1		PY2		PY3		PY4		PY5	
		CJR	Control Group	DiD	P-value	DiD	P-value	DiD	P-value	DiD	P-value	DiD	P-value
Age	20-64	3.2%	3.5%	-0.2	0.83	1.0	0.17	-0.4	0.64	0.0	0.97	0.1	0.88
	65-79	25.6%	28.3%	2.8	0.42	0.9	0.68	1.8	0.39	-2.9	0.19	1.1	0.51
	80+	71.1%	68.2%	-2.6	0.48	-2.0	0.44	-1.4	0.48	2.9	0.16	-1.2	0.49
Sex	Female	67.9%	66.5%	0.4	0.89	-2.6	0.26	-3.2	0.28	-1.4	0.46	-0.5	0.77
Race and ethnicity	White	87.6%	90.5%	0.7	0.70	-3.3	0.04	-3.1	0.03	-1.5	0.17	-0.4	0.74
	Black or African American	4.0%	5.2%	1.6	0.20	1.9	0.07	1.1	0.31	0.9	0.45	0.5	0.45
	Hispanic	5.4%	2.5%	-1.6	0.11	0.6	0.41	0.7	0.47	-0.2	0.81	-0.1	0.88
	Other[a]	3.0%	1.8%	-0.6	0.19	0.7	0.15	1.3	0.01	0.8	0.16	0.0	0.98
Medicaid	Eligible for Medicaid	24.0%	19.2%	2.7	0.26	2.7	0.20	3.8	0.06	0.4	0.86	1.0	0.54
Disability	Disability, no ESRD	11.8%	12.6%	1.2	0.62	2.9	0.09	0.8	0.64	-0.6	0.69	1.0	0.39
Health Status	HCC score	2.98	2.91	0.0	0.99	0.2	0.12	0.0	0.43	-0.1	0.98	0.0	0.89
	Obesity	5.0%	7.2%	0.8	0.65	1.7	0.17	0.2	0.91	1.3	0.42	3.2	0.02
	Diabetes	33.5%	30.0%	1.9	0.56	0.2	0.95	-1.5	0.56	-4.0	0.10	1.3	0.46
	Hypertension	78.8%	79.2%	0.1	0.96	2.6	0.13	3.3	0.04	0.9	0.67	0.6	0.71
	Dementia	34.9%	34.4%	-1.1	0.76	-0.5	0.80	8.7	<0.01	-0.4	0.83	1.5	0.37
	CHF	38.6%	37.0%	-1.7	0.93	2.1	0.04	0.9	0.63	-2.1	0.02	-1.8	0.57

Patient characteristics		Baseline Average		PY1		PY2		PY3		PY4		PY5	
		CJR	Control Group	DiD	P-value	DiD	P-value	DiD	P-value	DiD	P-value	DiD	P-value
Utilization in the six months prior to the anchor hospitalization	ACH stay	30.2%	28.9%	-3.5	0.24	1.1	0.64	1.7	0.42	-4.9	<0.01	-2.8	0.08
	HH use	32.1%	31.4%	-1.5	0.67	-1.7	0.53	-0.9	0.68	-4.1	0.07	-1.4	0.48
	IRF stay	3.5%	3.1%	-4.4	<0.01	0.2	0.83	-2.0	0.01	-0.9	0.18	-1.1	0.13
	SNF stay	15.9%	15.4%	-0.4	0.85	2.8	0.09	2.7	0.19	-0.8	0.50	-1.8	0.18
	Any prior care	56.6%	56.1%	-3.0	0.31	-1.6	0.62	-0.6	0.80	-3.3	0.07	-2.9	0.07

Source: CJR evaluation team analysis of Medicare claims and enrollment data for episodes initiated in 2012 through 2014 that ended between April 2012 and March 2015 (baseline) and episodes initiated during or after April 2016 that ended by September 2021 (intervention).

Notes: The estimates in this exhibit are the result of calculating the DiD of the unadjusted baseline and performance year averages for the CJR and control groups (net differences). DiD estimates that are significant at the 1%, 5%, or 10% significance level are indicated by red, orange, or yellow shaded cells, respectively.

Fracture is defined based on ICD codes for hip fracture provided by the CMMI on the CJR model website: https://innovation.cms.gov/Files/worksheets/cjr-icd10hipfracturecodes.xlsx.

ACH = acute care hospital, CHF = congestive heart failure, CI = confidence interval, CMMI = Center for Medicare & Medicaid Innovation, ESRD = end-stage renal disease, HCC = hierarchical condition category, HH = home health, ICD = International Classification of Diseases, IRF = inpatient rehabilitation facility, MCC = major complications or comorbidities, MS-DRG = Medicare Severity Diagnosis Related Group, pp = percentage point, PY = performance year, SNF = skilled nursing facility.

[a]   Other includes beneficiaries identified as "Asian," "American Indian/Alaska Native," or "Other."

Cumulative Results

Exhibit J-11: **Change in patient characteristics, fracture LEJRs with MCC (MS-DRG 469 or 522), PY1-5**

Patient characteristics		CJR		Control group		Net differences (pp)	Net differences % of baseline	p-value	90% CI
		Baseline Average (N=3,750)	Intervention Average (N=6,220)	Baseline Average (N=3,859)	Intervention Average (N=5,876)				
Age	20-64	3.2%	2.7%	3.5%	2.8%	0.2	4.7%	0.76	-0.7 to 1.0
	65-79	25.6%	29.4%	28.3%	31.6%	0.5	1.8%	0.72	-1.6 to 2.5
	80+	71.1%	67.9%	68.2%	65.6%	-0.6	-0.9%	0.64	-2.8 to 1.6
Sex	Female	67.9%	66.4%	66.5%	66.4%	-1.4	-2.0%	0.37	-3.9 to 1.1
Race and ethnicity	White[a]	87.6%	87.4%	90.5%	91.8%	-1.5	-1.7%	0.08	-2.9 to -0.1
	Black or African American[a]	4.0%	4.5%	5.2%	4.7%	1.1	26.3%	0.10	0.0 to 2.1
	Hispanic[a]	5.4%	5.0%	2.5%	2.1%	0.0	0.0%	1.00	-1.0 to 1.0
	Other[ab]	3.0%	3.1%	1.8%	1.4%	0.4	15.0%	0.15	-0.1 to 1.0
Medicaid	Eligible for Medicaid	24.0%	22.6%	19.2%	15.9%	1.8	7.4%	0.23	-0.7 to 4.2
Disability	Disability, no ESRD	11.8%	12.7%	12.6%	12.4%	1.0	8.5%	0.36	-0.8 to 2.8
Health Status	HCC score	3.0	3.1	2.9	3.1	0.0	-0.9%	0.56	-0.1 to 0.1
	Obesity	5.0%	10.3%	7.2%	10.6%	1.8	36.6%	0.07	0.1 to 3.5
	Diabetes	33.5%	32.5%	30.0%	29.5%	-0.4	-1.3%	0.79	-3.2 to 2.3
	Hypertension	78.8%	81.3%	79.2%	80.3%	1.4	1.8%	0.28	-0.7 to 3.5
	Dementia	34.9%	34.7%	34.4%	32.5%	1.7	4.9%	0.13	-0.2 to 3.6
	CHF	38.6%	35.6%	37.0%	34.7%	-0.7	-1.8%	0.63	-3.0 to 1.6
Utilization in the six months prior to the anchor hospitalization	ACH stay	30.2%	27.7%	28.9%	28.3%	-1.9	-6.3%	0.11	-3.8 to 0.0
	HH use	32.1%	29.2%	31.4%	30.3%	-1.9	-5.9%	0.30	-4.9 to 1.1
	IRF stay	3.5%	3.3%	3.1%	4.2%	-1.3	-36.1%	0.03	-2.2 to -0.3
	SNF stay	15.9%	14.9%	15.4%	14.3%	0.1	0.4%	0.95	-1.5 to 1.6
	Any prior care	56.6%	54.9%	56.1%	56.7%	-2.4	-4.2%	0.10	-4.7 to 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 initiated during or after April 2016 that ended by September 2021 (intervention).

LEWIN GROUP®

Notes: The estimates in this exhibit are the result of calculating the DiD of the unadjusted baseline and intervention averages for the CJR and control groups (net differences). DiD estimates that are significant at the 1%, 5%, or 10% significance level are indicated by red, orange, or yellow shaded cells, respectively.

Fracture is defined based on ICD codes for hip fracture provided by the CMMI on the CJR model website:
https://innovation.cms.gov/Files/worksheets/cjr-icd10hipfracturecodes.xlsx.

ACH = acute care hospital, CHF = congestive heart failure, CI = confidence interval, CMMI = Center for Medicare & Medicaid Innovation, ESRD = end-stage renal disease, HCC = hierarchical condition category, HH = home health, ICD = International Classification of Diseases, IRF = inpatient rehabilitation facility, MCC = major complications or comorbidities, MS-DRG = Medicare Severity Diagnosis Related Group, pp = percentage point, PY = performance year, SNF = skilled nursing facility.

[a] The number of episodes for these measures is lower because episodes were dropped to account for missing observations. These measures are based on 3,742 CJR baseline episodes, 6,220 CJR intervention episodes, 3,859 control group baseline episodes, and 5,876 control group intervention episodes.

[b] Other includes beneficiaries identified as "Asian," "American Indian/Alaska Native," or "Other."

Appendix K: Parallel Trends

Results of parallel trends tests

Exhibit K-1: **Linear and joint tests of parallel trends for payment, utilization, and quality metrics, mandatory CJR hospitals, baseline: all inpatient LEJR episodes, fracture episodes and elective episodes**

Domain	Measure	All LEJR		Fracture		Elective	
		Joint test	Linear test	Joint test	Linear test	Joint test	Linear test
Payments	Episode payments	p=0.64	p=0.88	p=0.98	p=0.77	p=0.54	p=0.72
	IRF payments	p=0.63	p=0.72	p=0.07	p=0.19	p=0.23	p=0.83
	SNF payments	p=0.55	p=0.94	p=0.31	p=0.35	p=0.27	p=0.89
	HH payments	p=0.01	p<0.01	p=0.49	p=0.66	p=0.02	p<0.01
	Readmission payments	p=0.41	p=0.28	p=0.96	p=0.99	p=0.37	p=0.25
	Part B payments	p=0.22	p=0.13	p=0.88	p=0.76	p=0.15	p=0.09
	30-day PEP payments	p<0.01	p=0.03	p=0.82	p=0.91	p<0.01	p=0.02
	Anchor payments	p=0.03	p=0.24	p=0.92	p=0.98	p=0.02	p=0.15
Utilization	First PAC IRF	p=0.60	p=0.75	p=0.11	p=0.26	p=0.19	p=0.86
	First PAC SNF	p=0.35	p=0.45	p=0.06	p=0.07	p=0.29	p=0.60
	First PAC HH	p=0.11	p=0.07	p=0.87	p=0.56	p=0.10	p=0.06
	First PAC home without HH	p=0.57	p=0.61	p=0.37	p=0.41	p=0.47	p=0.59
	Any HH use	p=0.19	p=0.09	p=0.76	p=1.00	p=0.24	p=0.10
	IRF days	p=0.60	p=0.49	p=0.56	p=0.41	p=0.80	p=0.75
	SNF days	p=0.69	p=0.18	p=0.74	p=0.20	p=0.66	p=0.25
	HH visits	p=0.42	p=0.20	p=0.76	p=0.71	p=0.40	p=0.19
	Outpatient PT/OT visits	p=0.07	p=0.04	p=0.19	p=0.20	p=0.11	p=0.04
Quality	Unplanned readmission rate	p=0.43	p=0.55	p=0.77	p=0.93	p=0.50	p=0.43
	ED use	p=0.71	p=0.67	p=0.24	p=0.12	p=0.81	p=0.86
	Mortality rate	p=0.87	p=0.86	p=0.99	p=0.86	p=0.74	p=0.50
	Complications[a]	NA	NA	NA	NA	p=0.92	p=0.72

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 are 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 are 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. We consider outcomes to *fail* parallel trends if we reject the null hypothesis of seemingly parallel trends for both tests at the 10% significance level. P-values of estimates that are significant at the 1%, 5%, or 10% significance level are indicated by red, orange, or yellow shaded cells, respectively.

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.

[a] The complications measure only applies to elective episodes.

Exhibit K-2:   **Linear and joint tests of parallel trends for payment metrics, mandatory CJR hospitals, baseline: fracture LEJR episodes and post-episode periods**

Measure	Period	Joint test	Linear test
Episode payments per days alive	Episode	p=0.88	p=0.96
	91-179 days PEP	p=0.85	p=0.88
	180-269 days PEP	p=0.09	p=0.70
	270-359 days PEP	p=0.22	p=0.20
IRF payments per days alive	Episode	p=0.06	p=0.12
	91-179 days PEP	p=0.87	p=0.48
	180-269 days PEP	p=0.01	p=0.03
	270-359 days PEP	p=0.67	p=0.30
SNF payments per days alive	Episode	p=0.52	p=0.38
	91-179 days PEP	p=0.96	p=0.96
	180-269 days PEP	p=0.34	p=0.69
	270-359 days PEP	p=0.21	p=0.95
HH payments per days alive	Episode	p=0.52	p=0.62
	91-179 days PEP	p=0.89	p=0.87
	180-269 days PEP	p=0.55	p=0.51
	270-359 days PEP	p=0.99	p=0.87
Readmission payments per days alive	Episode	p=0.91	p=0.90
	91-179 days PEP	p=0.83	p=0.59
	180-269 days PEP	p=0.06	p=0.68
	270-359 days PEP	p=0.27	p=0.24
Hospice payments per days alive	Episode	p=0.86	p=0.80
	91-179 days PEP	p=0.44	p=0.43
	180-269 days PEP	p=0.99	p=0.98
	270-359 days PEP	p=1.00	p=0.94
DME payments per days alive	Episode	p=0.46	p=0.70
	91-179 days PEP	p=0.49	p=0.35
	180-269 days PEP	p=0.89	p=0.90
	270-359 days PEP	p=0.99	p=0.70

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 are 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 are 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. We consider outcomes to *fail* parallel trends if we reject the null hypothesis of seemingly parallel trends for both tests at the 10% significance level. P-values of estimates that are significant at the 1%, 5%, or 10% significance level are indicated by red, orange, or yellow shaded cells, respectively.

DME = durable medical equipment, HH = home health, IRF = inpatient rehabilitation facility, LEJR = lower extremity joint replacement, PEP = post-episode period, SNF = skilled nursing facility.

Exhibit K-3: **Linear and joint tests of parallel trends for quality metrics, mandatory CJR hospitals, baseline: fracture LEJR episodes and post-episode periods**

Measure	Period	Joint test	Linear test
Unplanned readmission rate	Episode	p=0.35	p=0.58
	91-179 days PEP	p=0.67	p=0.83
	180-269 days PEP	p=0.65	p=0.35
	270-359 days PEP	p=0.58	p=0.56
ED use	Episode	p=0.29	p=0.16
	91-179 days PEP	p=0.40	p=0.30
	180-269 days PEP	p=0.91	p=0.89
	270-359 days PEP	p=0.28	p=0.41
Mortality rate	Episode	p=0.99	p=0.88
	91-179 days PEP	p=0.42	p=0.47
	180-269 days PEP	p<0.01	p=0.67
	270-359 days PEP	p=0.03	p=0.36
Pressure ulcers	Episode	p=0.53	p=0.38
	One year	p=0.54	p=0.25
Delirium	Episode	p=0.64	p=0.92
	One year	p=0.92	p=0.43

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 are 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 are 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. We consider outcomes to *fail* parallel trends if we reject the null hypothesis of seemingly parallel trends for both tests at the 10% significance level. P-values of estimates that are significant at the 1%, 5%, or 10% significance level are indicated by red, orange, or yellow shaded cells, respectively.

ED = emergency department, LEJR = lower extremity joint replacement, PEP = post-episode period.

Exhibit K-4:   Linear and joint tests of parallel trends for utilization metrics, mandatory CJR hospitals, baseline: fracture LEJR episodes and post-episode periods

Measure	Period	Joint test	Linear test
Any IRF use	Episode	p=0.07	p=0.19
	91-179 days PEP	p=0.66	p=0.34
	180-269 days PEP	p=0.17	p=0.15
	270-359 days PEP	p=0.30	p=0.11
Any NF use	Episode	p=0.40	p=0.09
	91-179 days PEP	p=0.65	p=0.46
	180-269 days PEP	p=0.08	p=0.04
	270-359 days PEP	p=0.02	p=0.02
Any SNF use	Episode	p=0.40	p=0.11
	91-179 days PEP	p=0.89	p=0.75
	180-269 days PEP	p=0.91	p=0.98
	270-359 days PEP	p=0.61	p=0.63
Any HH use	Episode	p=0.54	p=0.82
	91-179 days PEP	p=0.65	p=0.99
	180-269 days PEP	p=0.70	p=0.70
	270-359 days PEP	p=0.58	p=0.47
Any PT/OT use	Episode	p=0.50	p=0.11
	91-179 days PEP	p=0.65	p=0.39
	180-269 days PEP	p=0.18	p=0.20
	270-359 days PEP	p=0.28	p=0.89
Any hospice use	Episode	p=0.86	p=0.53
	91-179 days PEP	p=0.72	p=0.82
	180-269 days PEP	p=0.64	p=0.99
	270-359 days PEP	p=0.60	p=0.31
Days in Community	Episode	p=0.80	p=0.55
	91-179 days PEP	p=0.55	p=0.76
	180-269 days PEP	p=0.19	p=0.14
	270-359 days PEP	p=0.06	p=0.05

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 are 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 are 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. We consider outcomes to *fail* parallel trends if we reject the null hypothesis of seemingly parallel trends for both tests at the 10% significance level. P-values of estimates that are significant at the 1%, 5%, or 10% significance level are indicated by red, orange, or yellow shaded cells, respectively.

HH = home health, IRF = inpatient rehabilitation facility, LEJR = lower extremity joint replacement, NF = nursing facility, OT = occupational therapy, PEP = post-episode period, PT = physical therapy, SNF = skilled nursing facility.

Exhibit K-5:   **Linear and joint tests of parallel trends for DME metrics, mandatory CJR hospitals, baseline: fracture LEJR episodes and post-episode periods**

Measure	Period	Joint test	Linear test
Bathing	Episode	p=0.76	p=0.56
	One year	p=0.82	p=0.57
Lift	Episode	p=0.13	p=0.08
	One year	p=0.88	p=0.52
Mobility Aid	Episode	p=0.48	p=0.57
	One year	p=0.58	p=0.65
Beds	Episode	p=0.37	p=0.98
	91-179 days PEP	p=0.68	p=0.73
	180-269 days PEP	p=0.48	p=0.33
	270-359 days PEP	p=0.83	p=0.71
Wheelchair	Episode	p=0.88	p=0.60
	91-179 days PEP	p=0.77	p=0.75
	180-269 days PEP	p=0.89	p=0.90
	270-359 days PEP	p=0.98	p=0.97

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 are 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 are 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. We consider outcomes to *fail* parallel trends if we reject the null hypothesis of seemingly parallel trends for both tests at the 10% significance level. P-values of estimates that are significant at the 1%, 5%, or 10% significance level are indicated by red, orange, or yellow shaded cells, respectively.

DME = durable medical equipment, LEJR = lower extremity joint replacement, PEP = post-episode period.

Trends for outcomes that failed parallel trends

Exhibit K-6:   Risk-adjusted baseline trends for HH payments, all LEJR episodes at mandatory hospitals

Source: CJR evaluation team analysis of Medicare claims and enrollment data for episodes initiated in 2012 through 2014 that ended between April 2012 and March 2015.

Notes: Baseline trends were estimated from a risk-adjusted trend model using the baseline, interim, and intervention periods.

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

HH = home health, LEJR = lower extremity joint replacement.

Exhibit K-7: Risk-adjusted baseline trends for 30-day post-episode payments, all LEJR episodes at mandatory hospitals
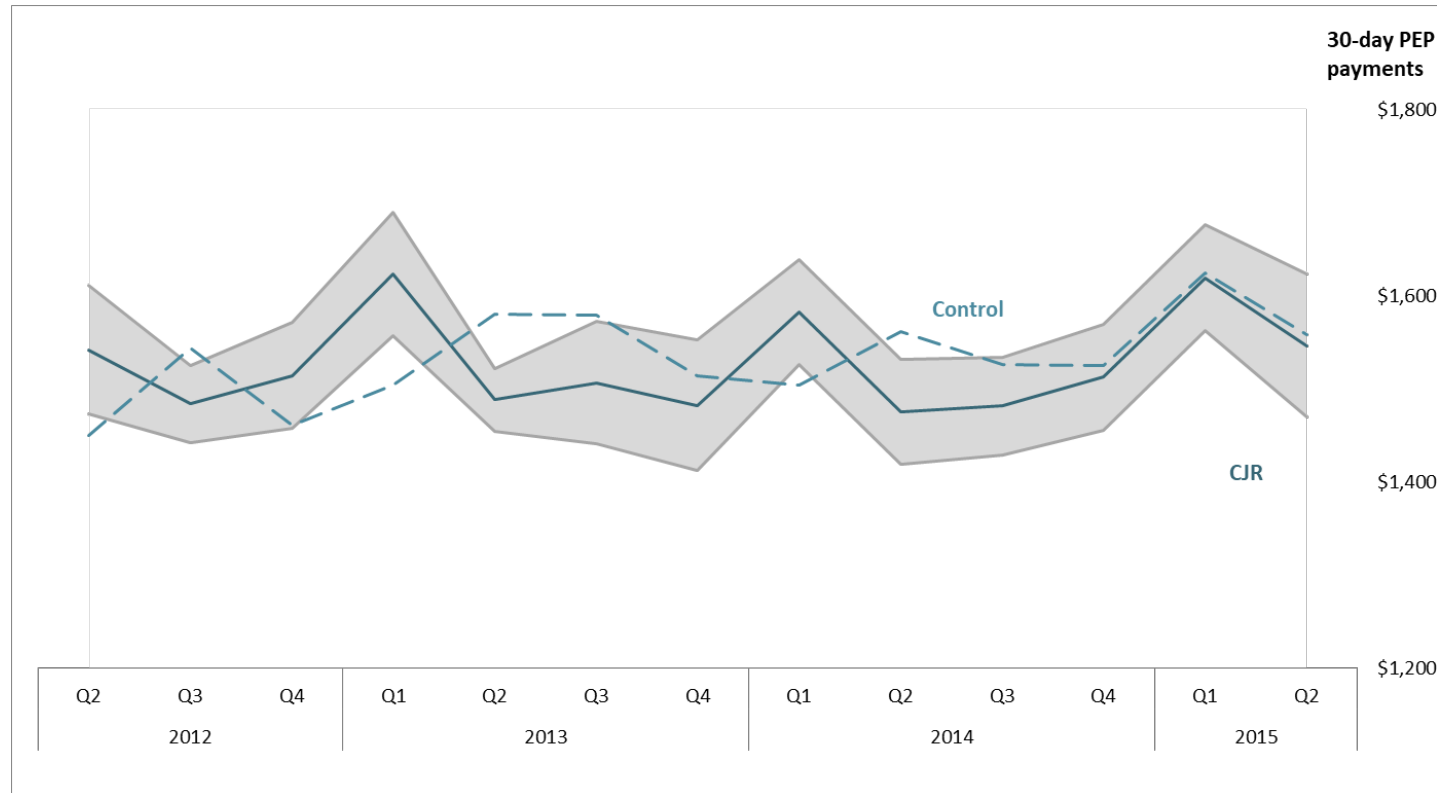
Source: CJR evaluation team analysis of Medicare claims and enrollment data for episodes initiated in 2012 through 2014 that ended between April 2012 and March 2015.

Notes: Baseline trends were estimated from a risk-adjusted trend model using the baseline, interim, and intervention periods.

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

LEJR = lower extremity joint replacement, PEP = post-episode period.

Exhibit K-8:   Risk-adjusted baseline trends for outpatient PT/OT visits, among PT/OT users, all LEJR episodes at mandatory hospitals
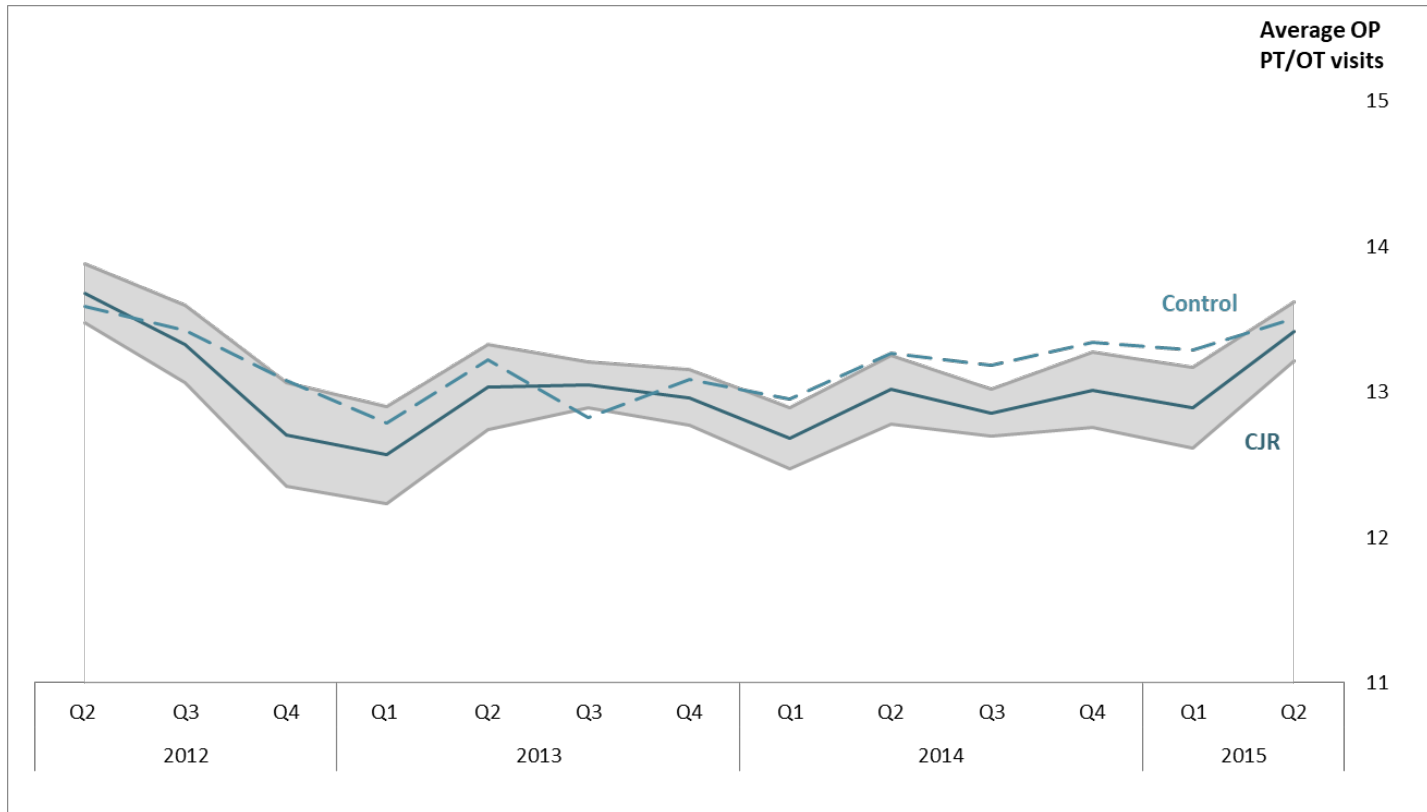
Source: CJR evaluation team analysis of Medicare claims and enrollment data for episodes initiated in 2012 through 2014 that ended between April 2012 and March 2015.

Notes: Baseline trends were estimated from a risk-adjusted trend model using the baseline, interim, and intervention periods.

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

LEJR = lower extremity joint replacement, OP = outpatient, OT = occupational therapy, PT = physical therapy.

Exhibit K-9: **Risk-adjusted baseline trends for first discharge to SNF, fracture LEJR episodes at mandatory hospitals**
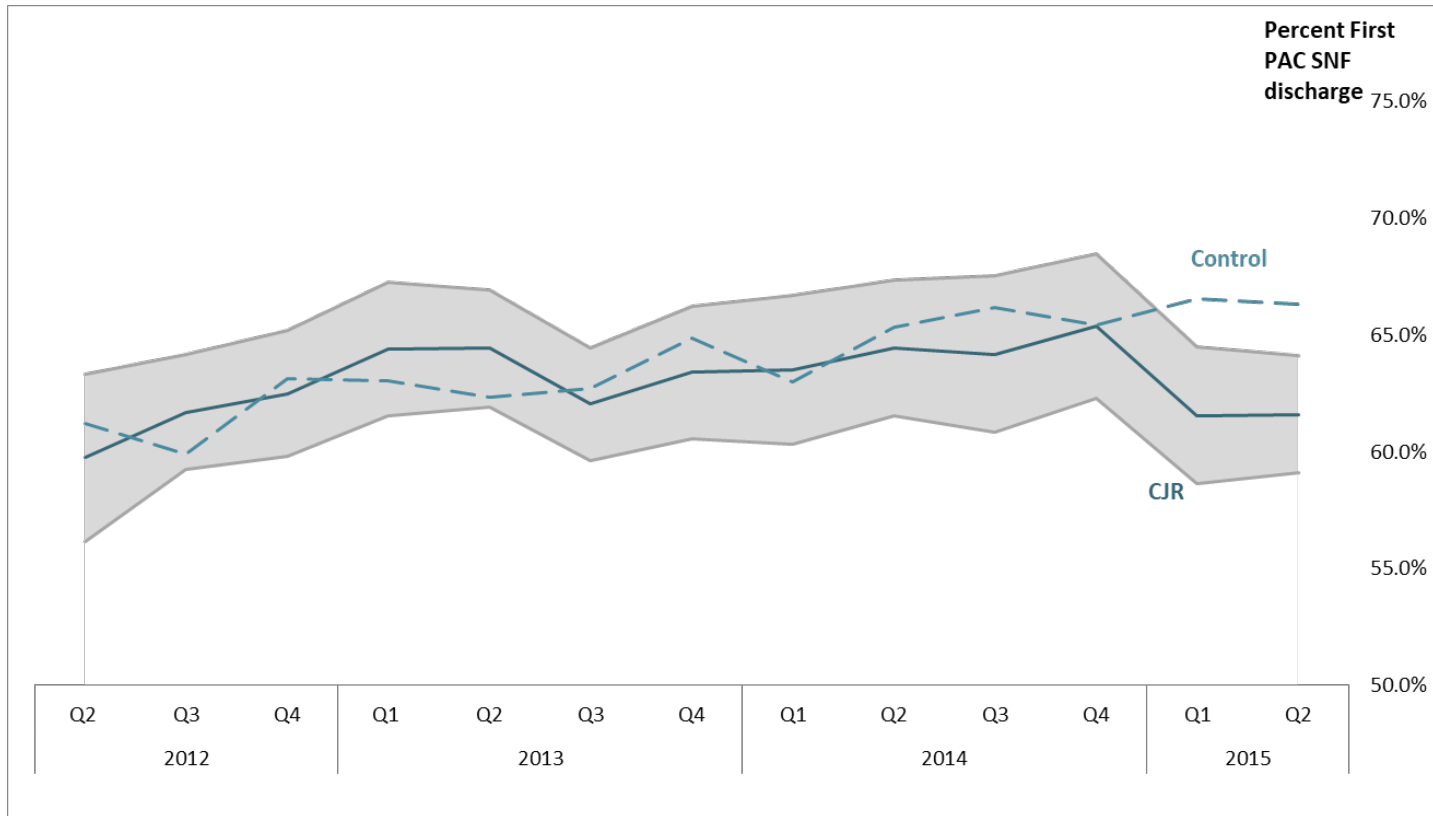
Source: CJR evaluation team analysis of Medicare claims and enrollment data for episodes initiated in 2012 through 2014 that ended between April 2012 and March 2015.

Notes: Baseline trends were estimated from a risk-adjusted trend model using the baseline, interim, and intervention periods.

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

LEJR = lower extremity joint replacement, PAC = post-acute care, SNF = skilled nursing facility.

Exhibit K-10: Risk-adjusted baseline trends for HH payments, elective LEJR episodes at mandatory hospitals
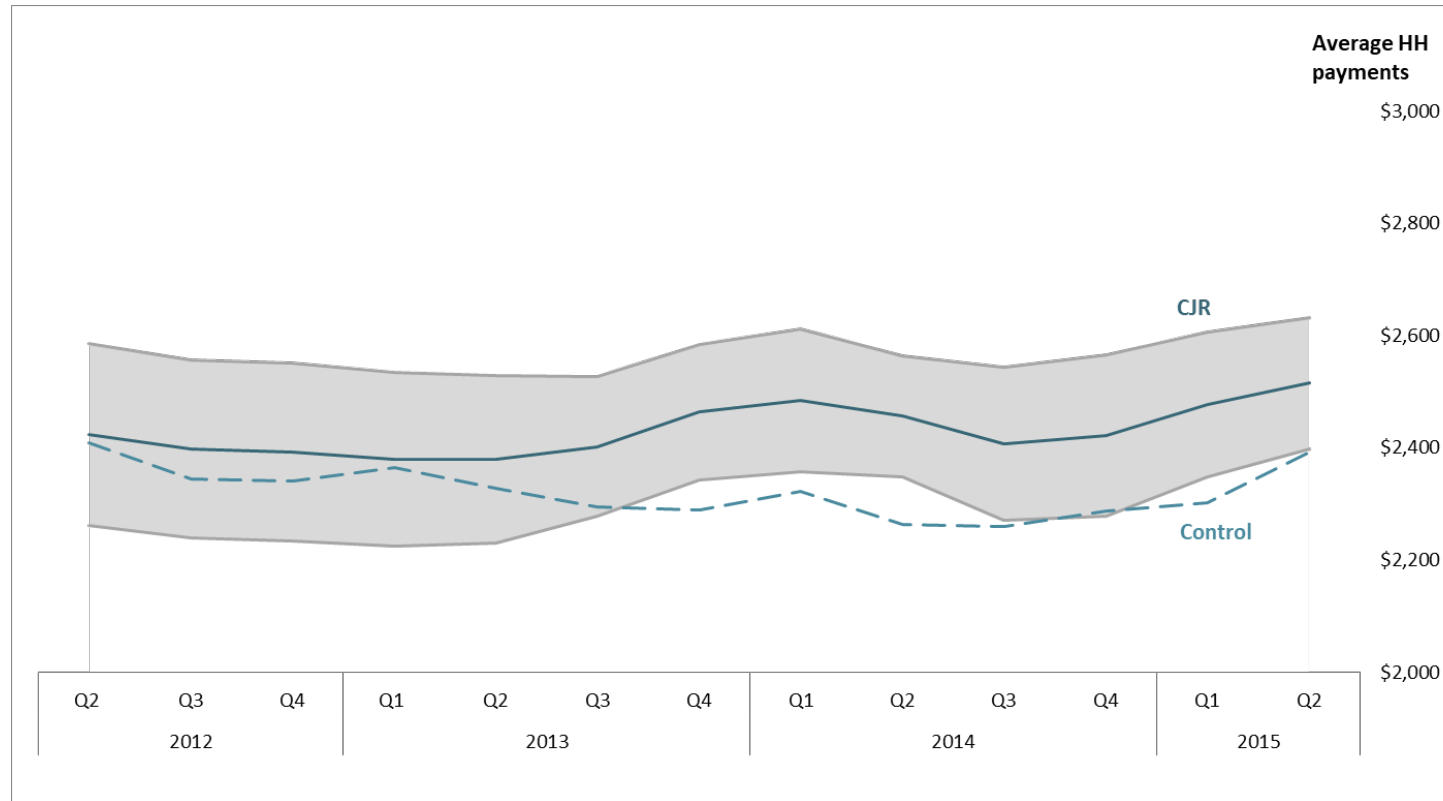
Source: CJR evaluation team analysis of Medicare claims and enrollment data for episodes initiated in 2012 through 2014 that ended between April 2012 and March 2015.

Notes: Baseline trends were estimated from a risk-adjusted trend model using the baseline, interim, and intervention periods.

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

HH = home health, LEJR = lower extremity joint replacement.

Exhibit K-11: Risk-adjusted baseline trends for 30-day post-episode payments, elective LEJR episodes at mandatory hospitals
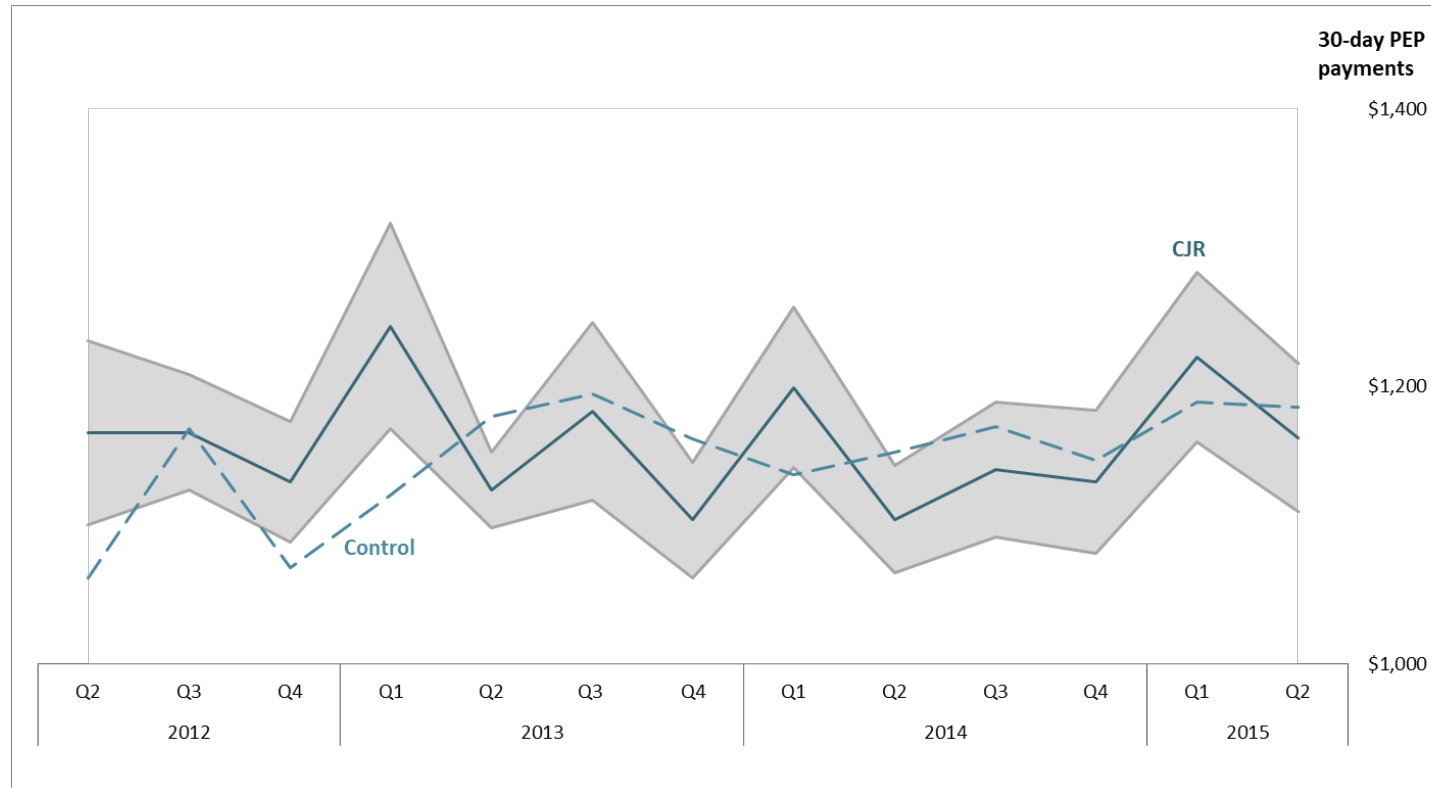
Source: CJR evaluation team analysis of Medicare claims and enrollment data for episodes initiated in 2012 through 2014 that ended between April 2012 and March 2015.

Notes: Baseline trends were estimated from a risk-adjusted trend model using the baseline, interim, and intervention periods.

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

LEJR = lower extremity joint replacement, PEP = post-episode period.

Exhibit K-12: Risk-adjusted baseline trends for IRF payments per days alive for days 180-269 PEP, fracture LEJR episodes at mandatory hospitals
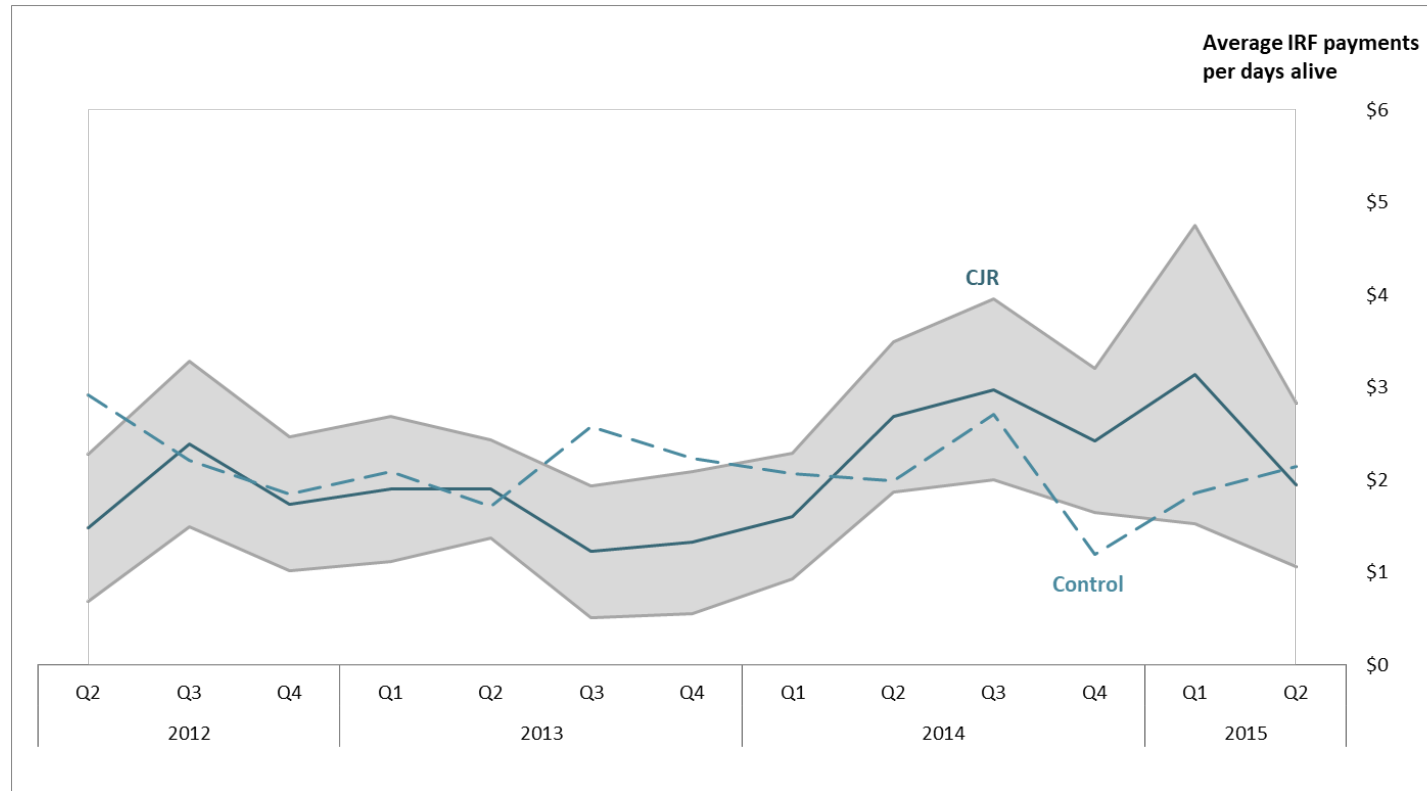
Source: CJR evaluation team analysis of Medicare claims and enrollment data for episodes initiated in 2012 through 2014 that ended between April 2012 and March 2015.

Notes: Baseline trends were estimated from a risk-adjusted trend model using the baseline, interim, and intervention periods.

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

IRF = inpatient rehabilitation facility, LEJR = lower extremity joint replacement, PEP = post-episode period.

LEWINGROUP®

Exhibit K-13: Risk-adjusted baseline trends for any prior NF use days 180-269 PEP, fracture LEJR episodes at mandatory hospitals
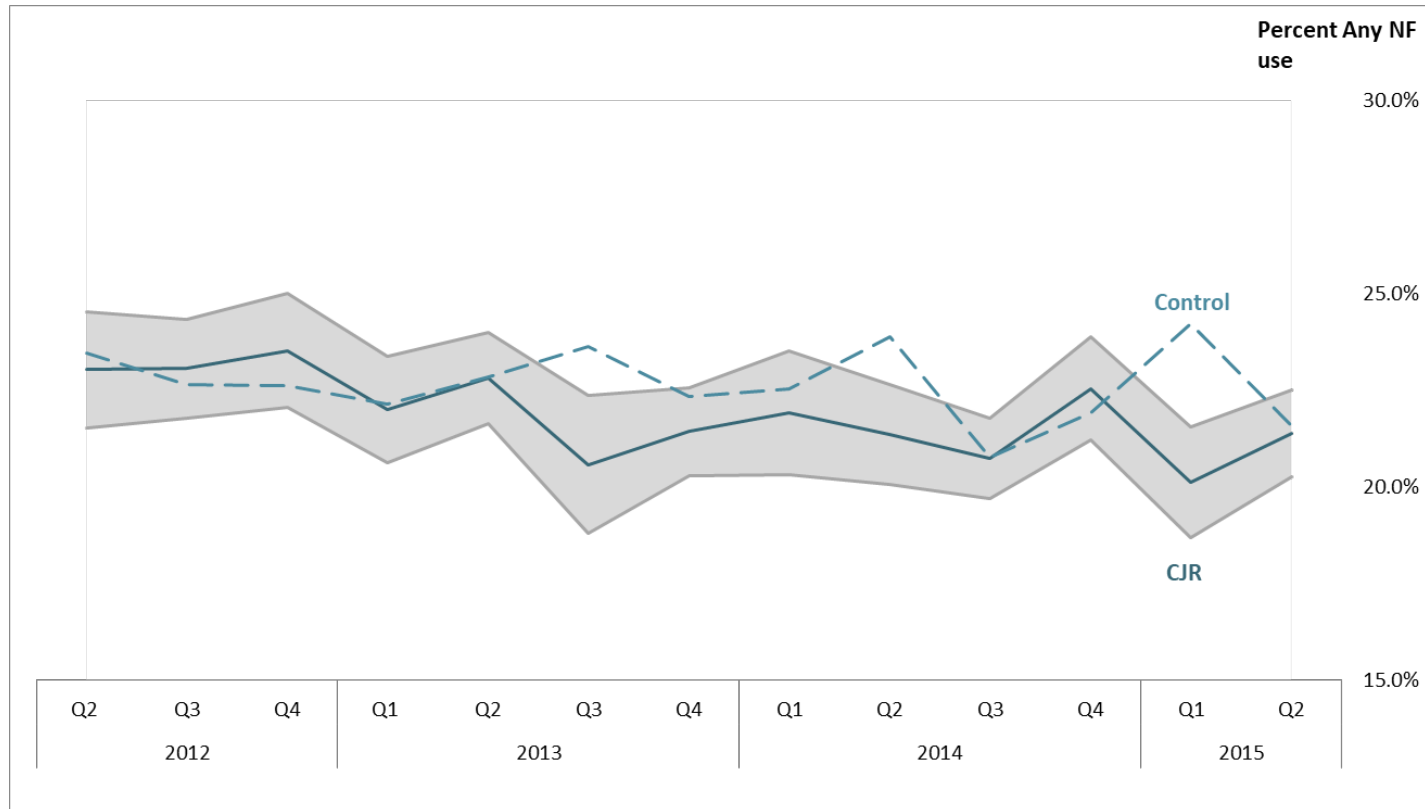
Source: CJR evaluation team analysis of Medicare claims and enrollment data for episodes initiated in 2012 through 2014 that ended between April 2012 and March 2015.

Notes: Baseline trends were estimated from a risk-adjusted trend model using the baseline, interim, and intervention periods.

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

LEJR = lower extremity joint replacement, NF = nursing facility, PEP = post-episode period.

Exhibit K-14: Risk-adjusted baseline trends for any prior NF use days 270-359 PEP, fracture LEJR episodes at mandatory hospitals
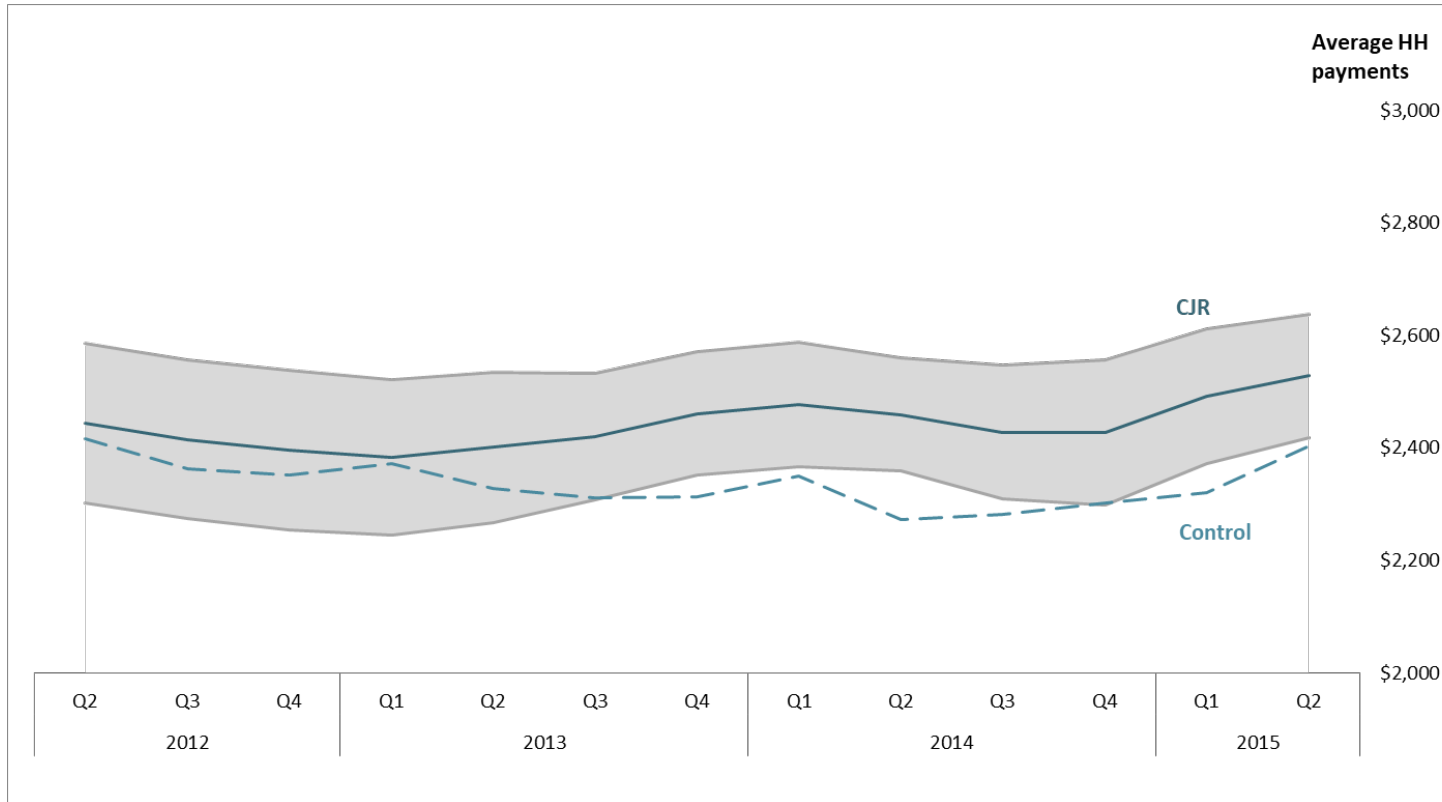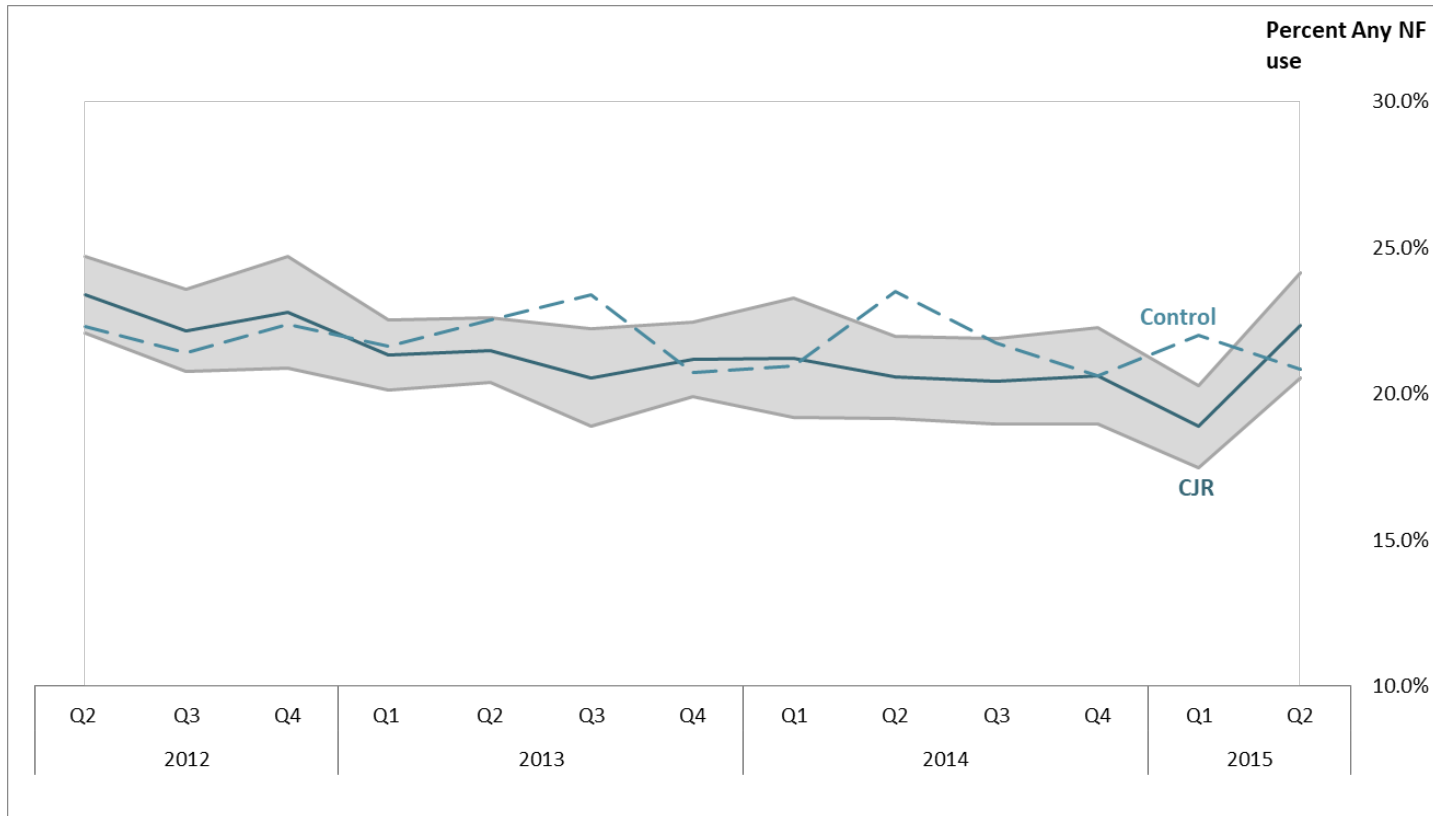

Source: CJR evaluation team analysis of Medicare claims and enrollment data for episodes initiated in 2012 through 2014 that ended between April 2012 and March 2015.

Notes: Baseline trends were estimated from a risk-adjusted trend model using the baseline, interim, and intervention periods.

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

LEJR = lower extremity joint replacement, NF = nursing facility, PEP = post-episode period.

Exhibit K-15: Risk-adjusted baseline trends for days in community days 270-359 PEP, fracture LEJR episodes at mandatory hospitals
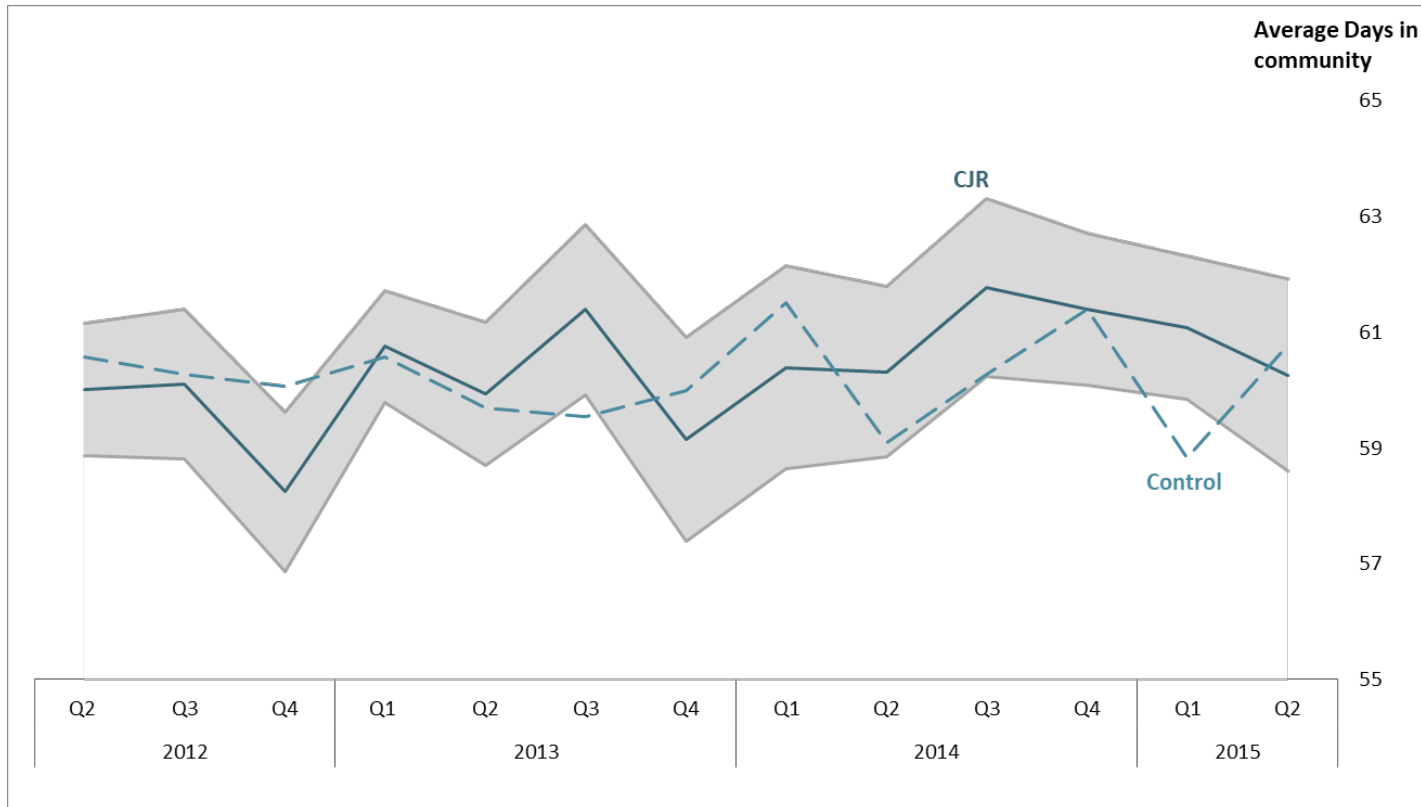
Source: CJR evaluation team analysis of Medicare claims and enrollment data for episodes initiated in 2012 through 2014 that ended between April 2012 and March 2015.

Notes: Baseline trends were estimated from a risk-adjusted trend model using the baseline, interim, and intervention periods.

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

LEJR = lower extremity joint replacement, PEP = post-episode period.

Appendix L: Payment, Utilization, and Quality Results for Fracture and Elective Episodes

Patient and Episode Characteristics

Exhibit L-1: Characteristics of fracture LEJR episodes compared to elective LEJR episodes, mandatory hospitals, baseline

| Domain	Measure	CJR fracture	CJR elective	p-value
Demographics	Mean age	82.5	73.2	<0.01
	Age 65-79	28.9%	69.6%	<0.01
	Age 80 or older	68.1%	21.7%	<0.01
	Female	74.0%	64.5%	<0.01
	Eligible for Medicaid	20.8%	13.2%	<0.01
	Disabled, not due to ESRD	3.1%	9.1%	<0.01
Health status	Mean HCC score	2.43	1.28	<0.01
	Cancer	16.8%	14.9%	<0.01
	Congestive heart failure	25.9%	12.8%	<0.01
	Dementia	30.1%	3.4%	<0.01
	Diabetes	29.6%	29.8%	0.52
	Hypertension	75.3%	75.4%	0.78
	Obesity	4.3%	17.7%	<0.01
Prior utilization	Any inpatient acute care hospitalizations	22.4%	11.4%	<0.01
	Any IRF stays	2.7%	1.1%	<0.01
	Any NF stays	16.7%	4.0%	<0.01
	Any SNF stays	12.2%	3.7%	<0.01
	Any HH use	26.1%	10.8%	<0.01
	Any prior care use	48.2%	26.8%	<0.01
Anchor hospitalization	Mean length of stay (days)	5.7	3.2	<0.01

Source: CJR evaluation team analysis of Medicare claims and enrollment data and Minimum Data Set (MDS) nursing facility assessment data for episodes initiated in 2012 through 2014 that ended between April 2012 and March 2015 (baseline).

Notes: ESRD = end stage renal disease, HCC = hierarchical condition category, HH = home health, IRF = inpatient rehabilitation facility, LEJR = lower extremity joint replacement, NF = nursing facility, SNF = skilled nursing facility.

Exhibit L-2: Setting of fracture beneficiaries from 1 to 360 days after anchor hospitalization, mandatory hospitals, baseline
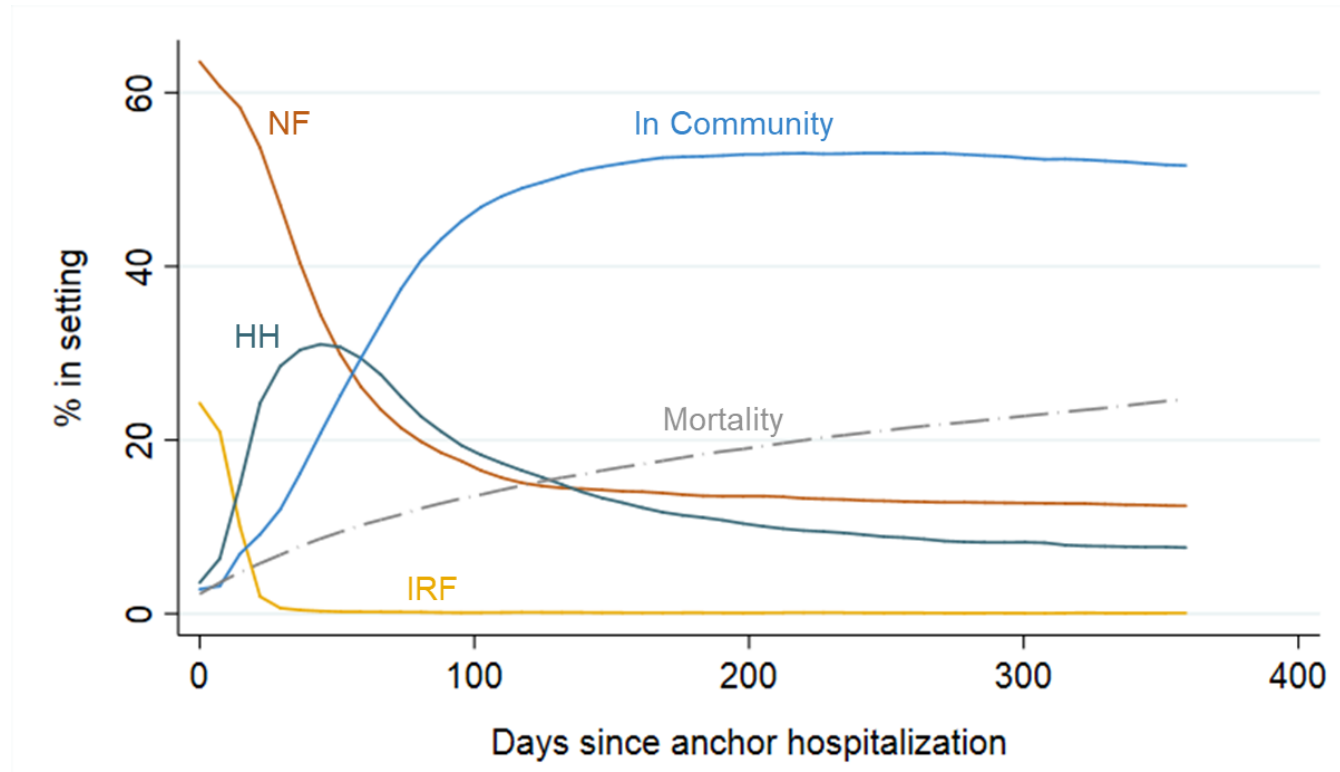
Source: CJR evaluation team analysis of Medicare claims and enrollment data and Minimum Data Set (MDS) nursing facility assessment data for hip fracture beneficiaries with LEJR procedures in 2012 through 2014 and that had 90-day episode periods ending between April 2012 and March 2015 (baseline). Data includes beneficiaries who died during the episode period.

Notes: HH = home health, IRF = inpatient rehabilitation facility, LEJR = lower extremity joint replacement, NF= nursing facility.

Fracture LEJRs

Performance-Year Specific Results – During the Episode

Exhibit L-3: Risk-adjusted claims-based difference-in-differences results for payment, utilization, and quality metrics during the episode, mandatory CJR hospitals, fracture LEJR episodes, by performance year for PY1-5

| Domain	Measure	Baseline risk-adjusted average		PY1		PY2		PY3		PY4		PY5.1		PY5.2	
		CJR	Control group	DiD	P-value	DiD	P-value	DiD	P-value	DiD	P-value	DiD	P-value	DiD	P-value
Payments	Episode payments	$47,376	$47,108	-$1,968	<0.01	-$1,678	<0.01	-$1,432	<0.01	-$2,311	<0.01	-$1,693	0.02	-$1,019	0.22
	IRF payments	$5,510	$5,156	-$681	0.03	-$359	0.24	-$515	0.22	-$1,150	0.01	-$636	0.21	-$283	0.64
	SNF payments	$16,569	$17,019	-$414	0.43	-$1,218	<0.01	-$1,013	0.07	-$608	0.27	-$720	0.24	-$383	0.60
	HH payments	$2,481	$2,411	-$14	0.84	$19	0.75	$65	0.27	$14	0.82	$70	0.37	$11	0.89
	Readmission payments	$2,607	$2,484	-$183	0.31	-$110	0.44	$5	0.98	-$240	0.07	-$5	0.98	$36	0.82
	Part B payments	$6,317	$6,160	-$293	0.01	$65	0.57	$93	0.27	-$124	0.34	-$196	0.06	-$19	0.89
	Anchor payments	$13,290	$13,288	-$72	0.27	$38	0.35	-$16	0.75	-$53	0.19	$7	0.89	$26	0.68
Utilization	First PAC IRF	28.4%	26.9%	-2.7	0.07	-2.7	0.07	-2.6	0.19	-5.8	<0.01	-3.8	0.12	-2.4	0.38
	First PAC SNF[a]	62.6%	63.1%	1.2	0.40	0.6	0.72	1.0	0.61	3.5	0.08	-1.4	0.40	-1.9	0.35
	First PAC HH	5.6%	6.4%	1.4	0.03	2.0	<0.01	1.4	0.10	1.5	0.08	4.5	<0.01	3.8	<0.01
	First PAC home without HH	3.3%	3.7%	0.1	0.93	0.2	0.63	0.3	0.50	0.8	0.04	0.6	0.13	0.5	0.26
	Any HH use	70.1%	67.9%	0.4	0.77	0.6	0.57	0.1	0.90	-0.5	0.64	0.2	0.87	-0.5	0.68
	IRF days	14.1	13.8	-0.6	0.07	-0.2	0.43	-0.3	0.29	-0.3	0.33	-0.3	0.27	-0.2	0.47
	SNF days	43.3	44.5	-0.9	0.35	-2.1	<0.01	-2.1	0.03	-1.6	0.07	-0.3	0.78	0.3	0.75
	HH visits	21.4	20.9	-0.5	0.36	-0.4	0.32	0.0	0.94	-0.2	0.76	-0.3	0.49	0.0	0.99
	Outpatient PT/OT visits	10.5	11.0	0.8	0.06	0.6	0.20	0.6	0.16	0.8	0.06	0.2	0.63	1.2	0.02
Quality	Unplanned readmission rate	19.9%	19.3%	-1.2	0.26	-0.8	0.28	0.2	0.80	-1.1	0.19	0.5	0.67	-0.1	0.94
	ED use	18.1%	17.1%	-0.6	0.65	-0.8	0.37	0.0	1.00	-0.6	0.41	-0.4	0.58	-0.6	0.63
	Mortality rate	12.7%	13.2%	-0.1	0.92	-0.2	0.82	-0.1	0.82	0.1	0.95	0.6	0.47	-0.8	0.38

LEWINGROUP®

Source: CJR evaluation team analysis of Medicare claims and enrollment data for episodes initiated in 2012 through 2014 that ended between April 2012 and March 2015 (baseline) and episodes initiated during or after April 2016 that ended by September 2021 (intervention).

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

The change in separate provider payments do not sum to the change in episode payments because separate models were estimated for episode payments and each component payment.

DiD = difference-in-differences, ED = emergency department, HH = home health, IRF = inpatient rehabilitation facility, LEJR = lower extremity joint replacement, OT = occupational therapy, PAC = post-acute care, PEP = post-episode period, PT = physical therapy, PY = performance year, SNF = skilled nursing facility.

[a]   We cannot be certain that there is no impact of the model because this outcome failed parallel trends tests (Appendix K). Parallel trends is an assumption that underlies our methodological approach. Please see Appendix C Section III.C.1.c for additional details.

Cumulative Results – During the Episode

Exhibit L-4: **Risk-adjusted claims-based difference-in-differences results for payment, utilization, and quality metrics during the episode, mandatory CJR hospitals, fracture LEJR episodes, PY1-5**

Domain	Measure	Intervention episodes (N)		CJR		Control group		DiD	DiD % of baseline	p-value	90% CI
		CJR	Control group	Baseline risk-adjusted average	Intervention risk-adjusted average	Baseline risk-adjusted average	Intervention risk-adjusted average				
Payments	Episode payments	34,102	31,966	$47,339	$45,705	$47,063	$47,139	-$1,710	-3.6%	<0.01	[-$2,487 to -$934]
	IRF payments	34,102	31,966	$5,598	$4,969	$5,263	$5,249	-$615	-11.0%	0.10	[-$1,228 to -$3]
	SNF payments	34,102	31,966	$16,398	$15,108	$16,796	$16,349	-$842	-5.1%	0.07	[-$1,594 to -$91]
	HH payments	34,102	31,966	$2,461	$2,638	$2,395	$2,533	$38	1.6%	0.42	[-$40 to $117]
	Readmission payments	34,102	31,966	$2,619	$2,548	$2,497	$2,497	-$72	-2.7%	0.43	[-$222 to $79]
	Part B payments	34,102	31,966	$6,281	$6,388	$6,129	$6,285	-$49	-0.8%	0.56	[-$189 to $90]
	Anchor payments	34,102	31,966	$13,392	$13,428	$13,388	$13,441	-$18	-0.1%	0.58	[-$72 to $36]
Utilization	First PAC IRF	34,101	31,966	28.3%	21.3%	26.8%	23.3%	-3.5	-12.4%	0.06	[-6.6 to -0.4]
	First PAC SNF[a]	34,101	31,966	62.5%	63.4%	62.7%	63.3%	0.4	0.6%	0.80	[-2.1 to 2.9]
	First PAC HH	34,101	31,966	5.9%	11.8%	6.7%	9.9%	2.7	45.7%	<0.01	[1.3 to 4.1]
	First PAC home without HH	34,101	31,966	3.3%	3.5%	3.7%	3.5%	0.4	12.8%	0.18	[-0.1 to 0.9]
	Any HH use	34,102	31,966	70.1%	72.3%	68.1%	70.1%	0.2	0.3%	0.79	[-1.0 to 1.4]
	IRF days	7,920	8,529	14.2	13.8	13.9	13.8	-0.3	-2.2%	0.17	[-0.7 to 0.1]
	SNF days	23,131	20,727	43.1	36.6	44.2	39.1	-1.4	-3.3%	0.03	[-2.5 to -0.3]
	HH visits	23,791	21,611	21.2	20.9	20.7	20.7	-0.2	-1.0%	0.62	[-0.9 to 0.5]
	Outpatient PT/OT visits	9,937	9,028	10.5	12.0	11.0	11.7	0.7	6.7%	0.04	[0.1 to 1.3]
	Days in the community[b]	52,079	50,659	19.8	23.9	20.5	23.0	1.6	8.1%	0.01	[0.6 to 2.6]

LEWINGROUP®

Domain	Measure	Intervention episodes (N)		CJR		Control group		DiD	DiD % of baseline	p-value	90% CI
		CJR	Control group	Baseline risk-adjusted average	Intervention risk-adjusted average	Baseline risk-adjusted average	Intervention risk-adjusted average				
Quality	Unplanned readmission rate	34,089	31,952	19.7%	18.9%	19.0%	18.6%	-0.4	-2.0%	0.54	[-1.5 to 0.7]
	ED use	34,089	31,952	18.0%	18.8%	17.0%	18.3%	-0.5	-2.7%	0.45	[-1.6 to 0.6]
	Mortality rate	38,422	36,094	13.1%	12.5%	13.5%	13.0%	-0.1	-0.4%	0.92	[-0.9 to 0.8]
	Complications[b]	52,070	48,170	12.4%	9.6%	12.1%	9.8%	-0.5	-4.3%	0.22	[-1.3 to 0.2]

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 initiated during or after April 2016 that ended by September 2021 (intervention).

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

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

The change in separate provider payments do not sum to the change in episode payments because separate models were estimated for episode payments and each component payment.

CI = confidence interval, DiD = difference-in-differences, ED = emergency department, HH = home health, IRF = inpatient rehabilitation facility, LEJR = lower extremity joint replacement, OT = occupational therapy, PAC = post-acute care, PEP = post-episode period, PT = physical therapy, PY = performance year, SNF = skilled nursing facility.

[a] We cannot be certain that there is no impact of the model because this outcome failed parallel trends tests (Appendix K). Parallel trends is an assumption that underlies our methodological approach. Please see Appendix C Section III.C.1.c for additional details.

[b] Includes beneficiaries who died during the measurement period. Also, the intervention period includes episodes through PY4.

Payment Outcome Results – Post-Episode

Exhibit L-5: Payment Outcomes: Risk-adjusted claims-based difference-in-differences results for episode and post-episode periods, mandatory CJR hospitals, fracture LEJR episodes, PY1-4

Measure	Period	Intervention episodes (N)		CJR		Control group		DiD	90-day equivalent DiD	DiD % of baseline	p-value	90% CI
		CJR	Control group	Baseline risk-adjusted average	Intervention risk-adjusted average	Baseline risk-adjusted average	Intervention risk-adjusted average					
Episode payments per days alive	Episode	52,079	50,659	$660	$638	$653	$656	-$24	-$2,189	-3.7%	<0.01	[-$38 to -$11]
	91-179 days PEP	45,621	44,133	$106	$107	$110	$112	-$1	-$63	-0.7%	0.75	[-$4 to $3]
	180-269 days PEP	42,912	41,323	$81	$85	$84	$85	$4	$348	4.8%	0.13	[$0 to $8]
	270-359 days PEP	41,025	39,397	$76	$80	$75	$76	$3	$232	3.4%	0.28	[-$1 to $7]
IRF payments per days alive	Episode	52,079	50,659	$58	$50	$55	$53	-$6	-$526	-10.1%	0.12	[-$12 to $0]
	91-179 days PEP	45,361	43,964	$2	$3	$2	$3	$0	$11	5.4%	0.76	[-$1 to $1]
	180-269 days PEP[a]	42,389	41,217	$2	$3	$2	$2	$0	$38	20.5%	0.20	[$0 to $1]
	270-359 days PEP	40,414	39,397	$2	$2	$2	$2	$1	$61	40.7%	0.02	[$0 to $1]
SNF payments per days alive	Episode	52,079	50,659	$183	$176	$186	$188	-$8	-$763	-4.6%	0.06	[-$16 to -$1]
	91-179 days PEP	45,621	44,133	$20	$19	$23	$20	$1	$80	4.4%	0.45	[-$1 to $3]
	180-269 days PEP	42,912	41,323	$12	$13	$13	$12	$1	$114	10.4%	0.17	[$0 to $3]
	270-359 days PEP	41,025	39,397	$13	$13	$13	$12	$1	$107	9.0%	0.15	[$0 to $3]

Measure	Period	Intervention episodes (N)		CJR		Control group		DiD	90-day equivalent DiD	DiD % of baseline	p-value	90% CI
		CJR	Control group	Baseline risk-adjusted average	Intervention risk-adjusted average	Baseline risk-adjusted average	Intervention risk-adjusted average					
HH payments per days alive	Episode	52,079	50,659	$25	$27	$24	$25	$1	$53	2.4%	0.17	[$0 to $1]
	91-179 days PEP	45,621	44,133	$13	$12	$13	$12	$0	$2	0.2%	0.96	[-$1 to $1]
	180-269 days PEP	42,912	41,323	$8	$8	$8	$7	$0	$17	2.3%	0.61	[$0 to $1]
	270-359 days PEP	41,025	39,397	$7	$7	$7	$6	$0	$20	3.2%	0.50	[$0 to $1]
Readmission payments per days alive	Episode	52,079	50,659	$44	$44	$45	$47	-$3	-$243	-6.1%	0.18	[-$6 to $1]
	91-179 days PEP	45,621	44,133	$27	$26	$26	$27	-$1	-$73	-3.1%	0.55	[-$3 to $1]
	180-269 days PEP	42,912	41,323	$21	$21	$21	$21	$1	$82	4.4%	0.45	[-$1 to $3]
	270-359 days PEP	41,025	39,397	$19	$20	$18	$18	$2	$160	9.4%	0.07	[$0 to $3]
Hospice payments per days alive	Episode	52,079	50,659	$2	$1	$2	$1	$0	$23	13.1%	0.03	[$0 to $0]
	91-179 days PEP	45,621	44,133	$2	$1	$2	$1	$0	$11	6.4%	0.23	[$0 to $0]
	180-269 days PEP	42,912	41,323	$2	$1	$2	$1	$0	$9	6.1%	0.33	[$0 to $0]
	270-359 days PEP	41,025	39,397	$2	$1	$2	$1	$0	$12	8.6%	0.26	[$0 to $0]
DME payments per days alive	Episode	52,079	50,659	$8	$9	$8	$10	$0	-$20	-3.0%	0.71	[-$1 to $1]
	91-179 days PEP	45,621	44,133	$8	$9	$7	$10	-$1	-$130	-18.4%	0.01	[-$2 to -$1]
	180-269 days PEP	42,912	41,323	$7	$8	$7	$9	$0	-$45	-7.1%	0.26	[-$1 to $0]
	270-359 days PEP	41,025	39,397	$7	$8	$7	$8	$0	-$37	-5.8%	0.41	[-$1 to $0]

Source: CJR evaluation team analysis of Medicare claims and enrollment data for hip fracture beneficiaries with LEJR procedures in 2012 through 2014 and that had 90-day episode periods ending between April 2012 and March 2015 (baseline) and LEJR procedures during or after April 2016 and that had 90-day episode periods ending by December 2019 (PY1-4 intervention period). Data includes beneficiaries who died during the given measurement period.

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

The 90-day equivalent DiD is calculated as the per day DiD estimate multiplied by 90. The relative change from CJR baseline is calculated as the DiD estimate as a percent of the CJR baseline level.

The change in separate provider payments do not sum to the change in episode payments because separate models were estimated for episode payments and each component payment.

CI = confidence interval, DiD = difference-in-differences, DME = durable medical equipment, HH = home health, IRF = inpatient rehabilitation facility, LEJR = lower extremity joint replacement, PEP = post-episode period, PY = performance year, SNF = skilled nursing facility.

[a] We cannot be certain that there is no impact of the model because this outcome failed parallel trends tests (Appendix K). Parallel trends is an assumption that underlies our methodological approach. Please see Appendix C Section III.C.1.c for additional details.

Utilization Outcome Results – Post-Episode

Exhibit L-6: Utilization Outcomes: Risk-adjusted claims-based difference-in-differences results for episode and post-episode periods, mandatory CJR hospitals, fracture LEJR episodes, PY1-4

Measure	Period	Intervention episodes (N)		CJR		Control group		DiD	DiD % of baseline	p-value	90% CI
		CJR	Control group	Baseline risk-adjusted average	Intervention risk-adjusted average	Baseline risk-adjusted average	Intervention risk-adjusted average				
Any IRF use	Episode	52,079	50,445	26.5%	20.3%	25.3%	22.0%	-2.9	-11.1%	0.06	[-5.5 to -0.4]
	91-179 days PEP	45,361	43,774	1.2%	1.3%	1.3%	1.3%	0.0	2.7%	0.87	[-0.3 to 0.4]
	180-269 days PEP	42,389	41,217	1.1%	1.2%	1.1%	1.0%	0.1	7.4%	0.55	[-0.1 to 0.3]
	270-359 days PEP	40,414	39,397	0.9%	1.0%	1.0%	0.9%	0.3	32.3%	0.02	[0.1 to 0.5]
Any NF use[b]	Episode	43,845	42,702	72.9%	73.3%	73.4%	73.7%	0.0	0.0%	1.00	[-1.9 to 1.9]
	91-179 days PEP	38,321	37,107	27.2%	24.4%	28.7%	24.8%	1.1	4.2%	0.07	[0.1 to 2.1]
	180-269 days PEP[a]	36,048	34,728	22.2%	20.4%	23.1%	20.3%	0.9	3.8%	0.12	[-0.1 to 1.8]
	270-359 days PEP[a]	34,448	33,112	21.5%	19.4%	22.4%	19.7%	0.6	2.7%	0.26	[-0.3 to 1.4]
Any SNF use	Episode	52,079	50,659	71.6%	72.1%	71.9%	72.0%	0.5	0.7%	0.68	[-1.5 to 2.5]
	91-179 days PEP	45,621	44,133	16.8%	13.7%	18.0%	14.3%	0.7	4.2%	0.35	[-0.5 to 1.9]
	180-269 days PEP	42,912	41,323	7.2%	6.9%	7.6%	6.7%	0.7	9.7%	0.06	[0.1 to 1.3]
	270-359 days PEP	41,025	39,397	7.3%	6.8%	7.2%	6.5%	0.1	1.4%	0.76	[-0.5 to 0.7]
Any HH use	Episode	52,079	50,659	62.7%	65.3%	60.9%	62.6%	0.9	1.4%	0.17	[-0.2 to 1.9]
	91-179 days PEP	45,621	44,133	40.4%	37.1%	39.1%	36.6%	-0.8	-2.1%	0.37	[-2.4 to 0.7]
	180-269 days PEP	42,912	41,323	29.6%	26.8%	28.2%	25.5%	-0.1	-0.5%	0.88	[-1.7 to 1.4]
	270-359 days PEP	41,025	39,397	26.8%	23.9%	25.6%	23.3%	-0.6	-2.2%	0.55	[-2.3 to 1.1]

Measure	Period	Intervention episodes (N)		CJR		Control group		DiD	DiD % of baseline	p-value	90% CI
		CJR	Control group	Baseline risk-adjusted average	Intervention risk-adjusted average	Baseline risk-adjusted average	Intervention risk-adjusted average				
Any PT/OT use	Episode	52,079	50,659	65.3%	70.0%	64.2%	68.1%	0.7	1.1%	0.26	[-0.3 to 1.8]
	91-179 days PEP	45,621	44,133	83.2%	85.7%	82.4%	84.6%	0.2	0.3%	0.68	[-0.7 to 1.1]
	180-269 days PEP	42,912	41,323	28.6%	31.6%	29.5%	31.4%	1.1	3.8%	0.20	[-0.3 to 2.5]
	270-359 days PEP	41,025	39,397	24.4%	26.6%	24.7%	26.9%	0.0	0.0%	1.00	[-1.7 to 1.7]
Any hospice use	Episode	52,079	50,659	8.5%	9.7%	8.8%	10.1%	-0.2	-2.7%	0.61	[-1 to 0.5]
	91-179 days PEP	45,621	44,133	7.4%	7.8%	7.3%	8.5%	-0.8	-11.0%	0.02	[-1.4 to -0.2]
	180-269 days PEP	42,912	41,323	6.4%	7.1%	6.5%	7.5%	-0.4	-6.5%	0.23	[-1 to 0.2]
	270-359 days PEP	41,025	39,397	6.1%	6.8%	6.1%	7.0%	-0.2	-3.6%	0.57	[-0.9 to 0.4]
Days in Community[b]	Episode	43,845	42,704	19.6	23.5	20.2	22.7	1.5	7.5%	0.02	[0.4 to 2.5]
	91-179 days PEP	38,321	37,109	51.8	54.7	51.4	54.0	0.2	0.5%	0.67	[-0.7 to 1.2]
	180-269 days PEP	36,048	34,730	58.3	60.6	58.0	60.9	-0.6	-1.1%	0.25	[-1.6 to 0.3]
	270-359 days PEP[a]	34,448	33,114	60.4	62.1	60.0	62.8	-1.1	-1.7%	0.08	[-2.0 to -0.1]

Source: CJR evaluation team analysis of Medicare claims and enrollment data and Minimum Data Set (MDS) nursing facility assessment data for hip fracture beneficiaries with LEJR procedures in 2012 through 2014 and that had 90-day episode periods ending between April 2012 and March 2015 (baseline) and LEJR procedures during or after April 2016 and that had 90-day episode periods ending by December 2019 (PY1-4 intervention period). Data includes beneficiaries who died during the given measurement period.

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

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

CI = confidence interval, DiD = difference-in-differences, HH = home health, IRF = inpatient rehabilitation facility, LEJR = lower extremity joint replacement, NF = nursing facility, OT = occupational therapy, PEP = post-episode period, PT = physical therapy, PY = performance year, SNF = skilled nursing facility.

[a] We cannot be certain that there is no impact of the model because this outcome failed parallel trends tests (Appendix K). Parallel trends is an assumption that underlies our methodological approach. Please see Appendix C Section III.C.1.c for additional details.

[b] This outcome only uses data from performance year 1 through performance year 3.

Quality Outcome Results – Post-Episode

Exhibit L-7: Quality Outcomes: Risk-adjusted claims-based difference-in-differences results for episode and post-episode periods, mandatory CJR hospitals, fracture LEJR episodes, PY1-4

Measure	Period	Intervention episodes (N)		CJR		Control group		DiD	DiD % of baseline	p-value	90% CI
		CJR	Control group	Baseline risk-adjusted average	Intervention risk-adjusted average	Baseline risk-adjusted average	Intervention risk-adjusted average				
Unplanned readmission rate	Episode	52,070	50,645	22.7%	22.0%	22.3%	22.2%	-0.6	-2.7%	0.39	[-1.8 to 0.5]
	91-179 days PEP	45,621	44,133	30.1%	27.3%	29.5%	27.2%	-0.5	-1.7%	0.50	[-1.7 to 0.7]
	180-269 days PEP	42,912	41,323	11.6%	10.7%	11.7%	10.9%	0.0	-0.1%	0.97	[-0.7 to 0.7]
	270-359 days PEP	41,025	39,397	10.9%	9.9%	11.4%	9.7%	0.7	6.0%	0.09	[0 to 1.3]
ED use	Episode	52,070	50,645	18.0%	19.2%	17.3%	18.9%	-0.2	-1.3%	0.75	[-1.4 to 0.9]
	91-179 days PEP	45,621	44,133	27.8%	28.9%	27.2%	28.7%	-0.5	-1.6%	0.57	[-1.8 to 0.9]
	180-269 days PEP	42,912	41,323	12.5%	12.6%	11.9%	12.5%	-0.5	-4.0%	0.38	[-1.4 to 0.4]
	270-359 days PEP	41,025	39,397	12.2%	11.9%	10.9%	11.3%	-0.7	-5.8%	0.14	[-1.5 to 0.1]
Mortality rate	Episode	52,079	50,659	12.8%	12.0%	13.5%	12.7%	-0.1	-0.6%	0.90	[-1 to 0.8]
	91-179 days PEP	45,621	44,133	6.2%	5.7%	6.5%	6.5%	-0.5	-8.0%	0.09	[-1 to 0]
	180-269 days PEP	42,912	41,323	4.5%	4.3%	4.9%	4.5%	0.2	3.7%	0.53	[-0.3 to 0.6]
	270-359 days PEP	41,025	39,397	3.9%	3.9%	4.2%	4.0%	0.2	5.8%	0.40	[-0.2 to 0.7]
Pressure ulcers	Episode	52,079	50,659	4.9%	6.1%	5.1%	5.8%	0.5	9.5%	0.21	[-0.1 to 1.1]
	One year	52,079	50,659	6.8%	7.9%	7.1%	7.5%	0.7	10.9%	0.10	[0 to 1.5]
Delirium	Episode	52,079	50,659	2.4%	3.8%	2.2%	3.9%	-0.2	-10.2%	0.42	[-0.7 to 0.2]
	One year	52,079	50,659	3.1%	4.8%	2.7%	4.7%	-0.2	-7.8%	0.49	[-0.8 to 0.3]

Source: CJR evaluation team analysis of Medicare claims and enrollment data for hip fracture beneficiaries with LEJR procedures in 2012 through 2014 and that had 90-day episode periods ending between April 2012 and March 2015 (baseline) and LEJR procedures during or after April 2016 and that had 90-day episode periods ending by December 2019 (PY1-4 intervention period).. Data includes beneficiaries who died during the given measurement period.

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

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

CI = confidence interval, DiD = difference-in-differences, ED = emergency department, LEJR = lower extremity joint replacement, PEP = post-episode period, PY = performance year.

DME Outcome Results – Post-Episode

Exhibit L-8: DME Outcomes: Risk-adjusted claims-based difference-in-differences results for episode and post-episode periods, mandatory CJR hospitals, fracture LEJR episodes, PY1-4

Measure	Period	Intervention episodes (N)		CJR		Control group		DiD	DiD % of baseline	p-value	90% CI
		CJR	Control group	Baseline risk-adjusted average	Intervention risk-adjusted average	Baseline risk-adjusted average	Intervention risk-adjusted average				
Bathing	Episode	52,079	50,303	20.2%	15.8%	19.2%	15.4%	-0.7	-3.4%	0.60	[-2.9 to 1.5]
	One year	52,079	50,303	21.9%	17.0%	21.0%	16.7%	-0.6	-2.9%	0.64	[-2.9 to 1.6]
Lift	Episode	52,079	50,303	0.9%	0.6%	0.8%	0.6%	-0.1	-13.0%	0.29	[-0.3 to 0.1]
	One year	52,079	50,659	1.7%	1.1%	1.4%	1.0%	-0.2	-11.6%	0.29	[-0.5 to 0.1]
Mobility Aid	Episode	52,079	50,659	26.1%	23.9%	25.8%	22.8%	0.8	3.1%	0.48	[-1.1 to 2.7]
	One year	52,079	50,659	28.4%	25.6%	27.9%	24.4%	0.8	2.8%	0.51	[-1.2 to 2.8]
Beds	Episode	52,079	50,659	7.4%	5.4%	6.6%	4.5%	0.0	-0.6%	0.94	[-1 to 1]
	91-179 days PEP	45,621	44,133	7.8%	5.5%	6.8%	4.9%	-0.4	-4.6%	0.54	[-1.3 to 0.6]
	180-269 days PEP	42,912	41,323	7.0%	5.1%	5.9%	4.3%	-0.3	-4.3%	0.53	[-1.1 to 0.5]
	270-359 days PEP	41,025	39,397	6.4%	4.8%	5.4%	4.0%	-0.3	-4.3%	0.42	[-0.8 to 0.3]
Wheelchair	Episode	52,079	50,659	18.6%	16.2%	16.0%	13.9%	-0.2	-0.9%	0.89	[-2.1 to 1.8]
	91-179 days PEP	45,621	44,133	19.6%	17.3%	17.2%	15.1%	-0.2	-1.0%	0.84	[-1.8 to 1.4]
	180-269 days PEP	42,912	41,323	18.4%	16.5%	16.0%	14.0%	0.0	0.1%	0.97	[-1.2 to 1.3]
	270-359 days PEP	41,025	39,397	16.9%	15.6%	14.6%	13.1%	0.3	1.7%	0.70	[-0.9 to 1.5]

Source: CJR evaluation team analysis of Medicare claims and enrollment data for hip fracture beneficiaries with LEJR procedures in 2012 through 2014 and that had 90-day episode periods ending between April 2012 and March 2015 (baseline) and LEJR procedures during or after April 2016 and that had 90-day episode periods ending by December 2019 (PY1-4 intervention period). Data includes beneficiaries who died during the given measurement period.

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

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

LEWINGROUP®

Elective LEJRs

Performance-Year Specific Results – During the Episode

Exhibit L-9: Risk-adjusted claims-based difference-in-differences results for payment, utilization, and quality metrics during the episode, mandatory CJR hospitals, elective LEJR episodes, by performance year for PY1-5

Domain	Measure	Baseline risk-adjusted average		PY1		PY2		PY3		PY4		PY5.1		PY5.2	
		CJR	Control group	DiD	P-value	DiD	P-value	DiD	P-value	DiD	P-value	DiD	P-value	DiD	P-value
Payments	Episode payments	$26,054	$25,417	-$1,387	<0.01	-$1,644	<0.01	-$1,375	<0.01	-$1,041	0.04	-$966	0.09	-$1,059	0.10
	IRF payments	$1,673	$1,594	-$531	<0.01	-$495	<0.01	-$503	<0.01	-$606	<0.01	-$485	0.03	-$387	0.15
	SNF payments	$4,210	$4,171	-$716	<0.01	-$882	<0.01	-$743	<0.01	-$423	0.09	-$252	0.35	-$415	0.17
	HH payments[a]	$2,412	$2,307	$137	0.16	$38	0.75	$77	0.65	$34	0.86	$92	0.65	$312	0.14
	Readmission payments	$994	$869	-$100	0.15	-$153	0.10	-$171	0.07	-$146	0.08	-$118	0.22	-$184	0.10
	Part B payments	$4,808	$4,608	-$142	0.02	-$132	0.06	-$53	0.42	$50	0.54	-$84	0.38	-$73	0.53
	Anchor payments	$12,008	$12,010	$23	0.72	$31	0.60	$113	0.03	$193	<0.01	$76	0.04	$15	0.69
Utilization	First PAC IRF	11.7%	11.1%	-3.7	<0.01	-3.5	<0.01	-3.6	<0.01	-4.5	<0.01	-3.2	0.05	-2.5	0.17
	First PAC SNF	36.9%	37.6%	-2.8	0.02	-4.2	<0.01	-3.2	0.13	0.7	0.78	-0.5	0.85	-3.1	0.30
	First PAC HH	42.1%	38.5%	8.9	<0.01	8.5	0.01	8.9	0.04	4.8	0.33	7.1	0.21	13.3	0.03
	First PAC home without HH	9.2%	12.8%	-2.3	0.29	-0.8	0.74	-2.1	0.56	-0.9	0.79	-3.4	0.40	-7.7	0.05
	Any HH use	75.3%	72.6%	5.9	0.03	4.9	0.13	5.6	0.27	3.6	0.54	6.3	0.34	13.0	0.06
	IRF days	10.4	10.2	0.1	0.73	0.1	0.85	-0.1	0.64	0.3	0.32	-0.4	0.13	-0.4	0.17
	SNF days	20.6	20.4	-2.3	<0.01	-2.8	<0.01	-2.6	<0.01	-2.8	<0.01	-1.5	0.02	-0.2	0.76
	HH visits	16.3	16.2	-0.7	0.07	-1.3	<0.01	-1.3	0.01	-1.1	0.06	-0.9	0.12	-0.6	0.36
	Outpatient PT/OT visits	13.2	13.4	0.3	0.20	0.5	0.15	0.3	0.23	0.4	0.22	0.4	0.19	0.0	0.92

Domain	Measure	Baseline risk-adjusted average		PY1		PY2		PY3		PY4		PY5.1		PY5.2	
		CJR	Control group	DiD	P-value	DiD	P-value	DiD	P-value	DiD	P-value	DiD	P-value	DiD	P-value
Quality	Unplanned readmission rate	7.3%	7.1%	-0.1	0.69	-0.3	0.27	-0.2	0.47	0.0	0.99	0.1	0.64	0.0	0.95
	ED use	12.3%	11.9%	0.1	0.87	-0.2	0.65	0.6	0.06	0.7	0.07	0.1	0.72	0.0	0.98
	Mortality rate	0.6%	0.6%	0.0	0.69	0.0	0.91	0.1	0.08	0.0	0.42	0.0	0.36	0.0	0.95
	Complications	3.3%	3.1%	0.2	0.32	-0.4	<0.01	-0.3	0.10	-0.1	0.52	-0.1	0.32	-0.2	0.26

Source: CJR evaluation team analysis of Medicare claims and enrollment data for episodes initiated in 2012 through 2014 that ended between April 2012 and March 2015 (baseline) and episodes initiated during or after April 2016 that ended by September 2021 (intervention).

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

The change in separate provider payments do not sum to the change in episode payments because separate models were estimated for episode payments and each component payment.

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

DiD = difference-in-differences, ED = emergency department, HH = home health, IRF = inpatient rehabilitation facility, LEJR = lower extremity joint replacement, OT = occupational therapy, PAC = post-acute care, PEP = post-episode period, PT = physical therapy, PY = performance year, SNF = skilled nursing facility, TKA = total knee arthroplasty.

[a] We cannot be certain that there is no impact of the model because this outcome failed parallel trends tests (Appendix K). Parallel trends is an assumption that underlies our methodological approach. Please see Appendix C Section III.C.1.c for additional details.

LEWIN GROUP®

Cumulative Results – During the Episode

Exhibit L-10: Risk-adjusted claims-based difference-in-differences results for payment, utilization, and quality metrics during the episode, mandatory CJR hospitals, elective LEJR episodes, PY1-5

Domain	Measure	Intervention episodes (N)		CJR		Control group		DiD	DiD % of baseline	p-value	90% CI
		CJR	Control group	Baseline risk-adjusted average	Intervention risk-adjusted average	Baseline risk-adjusted average	Intervention risk-adjusted average				
Payments	Episode payments	171,791	242,874	$25,984	$22,835	$25,345	$23,639	-$1,444	-5.6%	<0.01	[-$2,186 to -$701]
	IRF payments	171,791	242,874	$1,662	$664	$1,582	$1,110	-$525	-31.6%	<0.01	[-$816 to -$235]
	SNF payments	171,791	242,874	$4,154	$2,242	$4,110	$2,906	-$707	-17.0%	<0.01	[-$1,070 to -$345]
	HH payments[a]	171,791	242,874	$2,395	$2,394	$2,290	$2,216	$73	3.1%	0.64	[-$182 to $329]
	Readmission payments	171,791	242,874	$995	$824	$871	$850	-$150	-15.1%	0.08	[-$288 to -$11]
	Part B payments	171,791	242,874	$4,780	$4,599	$4,581	$4,490	-$90	-1.9%	0.18	[-$203 to $22]
	Anchor payments	171,791	242,874	$12,025	$12,015	$12,027	$11,958	$59	0.5%	0.19	[-$15 to $132]
Utilization	First PAC IRF	171,791	242,874	11.5%	3.3%	10.9%	6.5%	-3.8	-32.8%	<0.01	[-5.9 to -1.6]
	First PAC SNF	171,791	242,874	36.5%	21.3%	37.1%	25.1%	-3.2	-8.7%	0.10	[-6.4 to 0.0]
	First PAC HH	171,791	242,874	42.7%	57.5%	39.0%	45.3%	8.5	19.9%	0.04	[1.7 to 15.3]
	First PAC home without HH	171,791	242,874	9.4%	18.0%	13.0%	23.1%	-1.5	-16.2%	0.60	[-6.3 to 3.3]
	Any HH use	171,791	242,874	75.2%	76.3%	72.5%	68.1%	5.5	7.3%	0.25	[-2.3 to 13.3]
	IRF days	6,803	11,950	10.4	10.9	10.2	10.7	0.0	0.0%	1.00	[-0.3 to 0.3]
	SNF days	38,132	42,813	20.6	16.4	20.4	18.6	-2.4	-11.7%	<0.01	[-3.6 to -1.2]
	HH visits	126,451	157,376	16.2	14.6	16.1	15.6	-1.2	-7.2%	0.02	[-2.0 to -0.4]
	Outpatient PT/OT visits	127,766	192,146	13.1	14.8	13.3	14.6	0.4	2.9%	0.16	[-0.1 to 0.8]

LEWINGROUP®

| Domain | Measure | Intervention episodes (N) | | CJR | | Control group | | | | | |
		CJR	Control group	Baseline risk-adjusted average	Intervention risk-adjusted average	Baseline risk-adjusted average	Intervention risk-adjusted average	DiD	DiD % of baseline	p-value	90% CI
Quality	Unplanned readmission rate	171,734	242,836	7.3%	6.9%	7.1%	6.9%	-0.2	-2.2%	0.30	[-0.4 to 0.1]
	ED use	171,734	242,836	12.2%	12.9%	11.8%	12.3%	0.2	1.7%	0.47	[-0.3 to 0.7]
	Mortality rate	172,576	243,813	0.6%	0.5%	0.6%	0.5%	0.0	3.6%	0.56	[-0.0 to 0.1]
	Complications	171,734	242,836	3.3%	2.5%	3.1%	2.5%	-0.2	-7.4%	0.03	[-0.4 to -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 April 2016 that ended by September 2021 (intervention).

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

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

The change in separate provider payments do not sum to the change in episode payments because separate models were estimated for episode payments and each component payment.

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

CI = confidence interval, DiD = difference-in-differences, ED = emergency department, HH = home health, IRF = inpatient rehabilitation facility, LEJR = lower extremity joint replacement, OT = occupational therapy, PAC = post-acute care, PEP = post-episode period, PT = physical therapy, PY = performance year, SNF = skilled nursing facility, TKA = total knee arthroplasty.

[a] We cannot be certain that there is no impact of the model because this outcome failed parallel trends tests (Appendix K). Parallel trends is an assumption that underlies our methodological approach. Please see Appendix C Section III.C.1.c for additional details.

Appendix M: Additional Findings Related to the Impact of the CJR Model on Hip Fracture Patients

This supplemental appendix to the fracture chapter provides findings about the most common care pathways for CJR fracture patients and the impact of the CJR model on the long-term health trajectory of fracture patients post-episode.

Characteristics and care pathways of CJR fracture LEJR during baseline

Exhibit M-1: Most common PAC pathways for CJR fracture patients at baseline
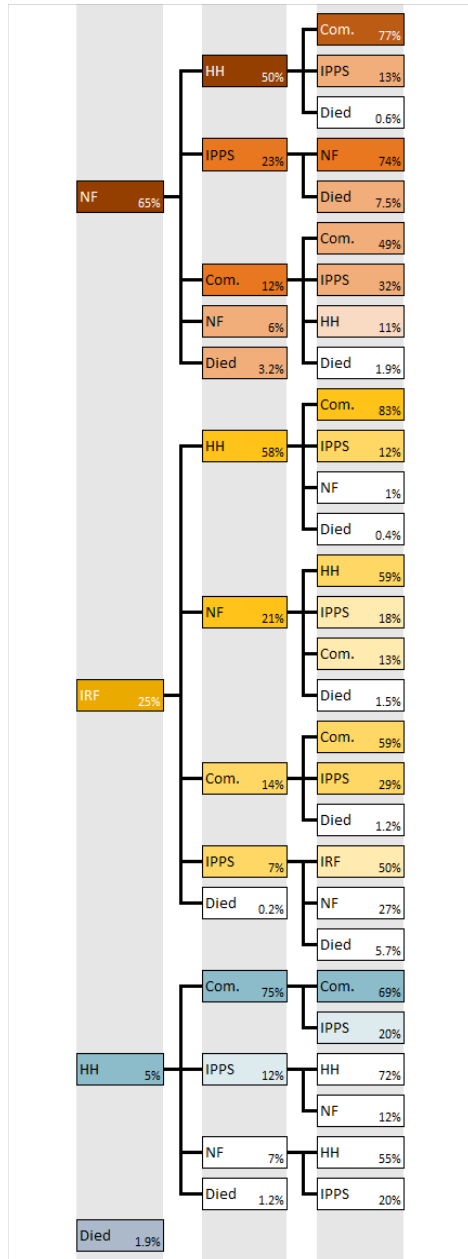
We studied care pathways for fracture patients to understand their care utilization from hospital discharge after the LEJR to up to the following 360 days. These pathways consisted of unique sequences of different care settings (e.g., NF, IRF, HH, inpatient stay, hospice). On days when a patient did not receive care in an institutional setting, hospice, or HH, we categorized them as "in the community." While lack of institutional, hospice, or HH care could indicate care was not required, our data precludes us from knowing the health status of patients in the community. See Appendix C Section VI for additional details relating to how we measured care pathways.

CJR fracture patients followed numerous care pathways after their LEJR surgery. Most commonly, fracture patients went to a nursing facility after their surgery, then home with HH, and then were in the community (Exhibit M-1). A few fracture patients were discharged directly home with HH after their LEJR (5%), but most patients received HH after being released from a nursing facility (50%) or IRF (58%).

Exhibit M-2 illustrates the median amount of time fracture patients spent in each setting for the most common pathways. For the most common pathway, during the baseline, patients spent an average of 32 days in a nursing facility and then 36 days at home receiving HH care before returning to the community. In general, patients discharged from the hospital to an IRF went home with or without home health services earlier than patients discharged from the hospital to a nursing facility.

During the baseline, by the end of the episode, 43% of CJR fracture patients were in the community or at home without HH care, 21% were receiving HH care, 20% were in an institutional setting, 3% were receiving hospice care, and 13% had died (Appendix L, Exhibit L-2).

Source: CJR evaluation team analysis of Medicare claims and enrollment data and MDS assessment data for fracture episodes from January 2012 through March 2015 (baseline).

Notes: The most common pathways are depicted, so percentages do not sum to 100%. Care pathways are measured up to 360 days after LEJR discharge. Darker colors indicate more common pathways (greater volume).

LEWINGROUP®

Exhibit M-2: Median number of days in each setting after discharge from the hospital, CJR fracture patients
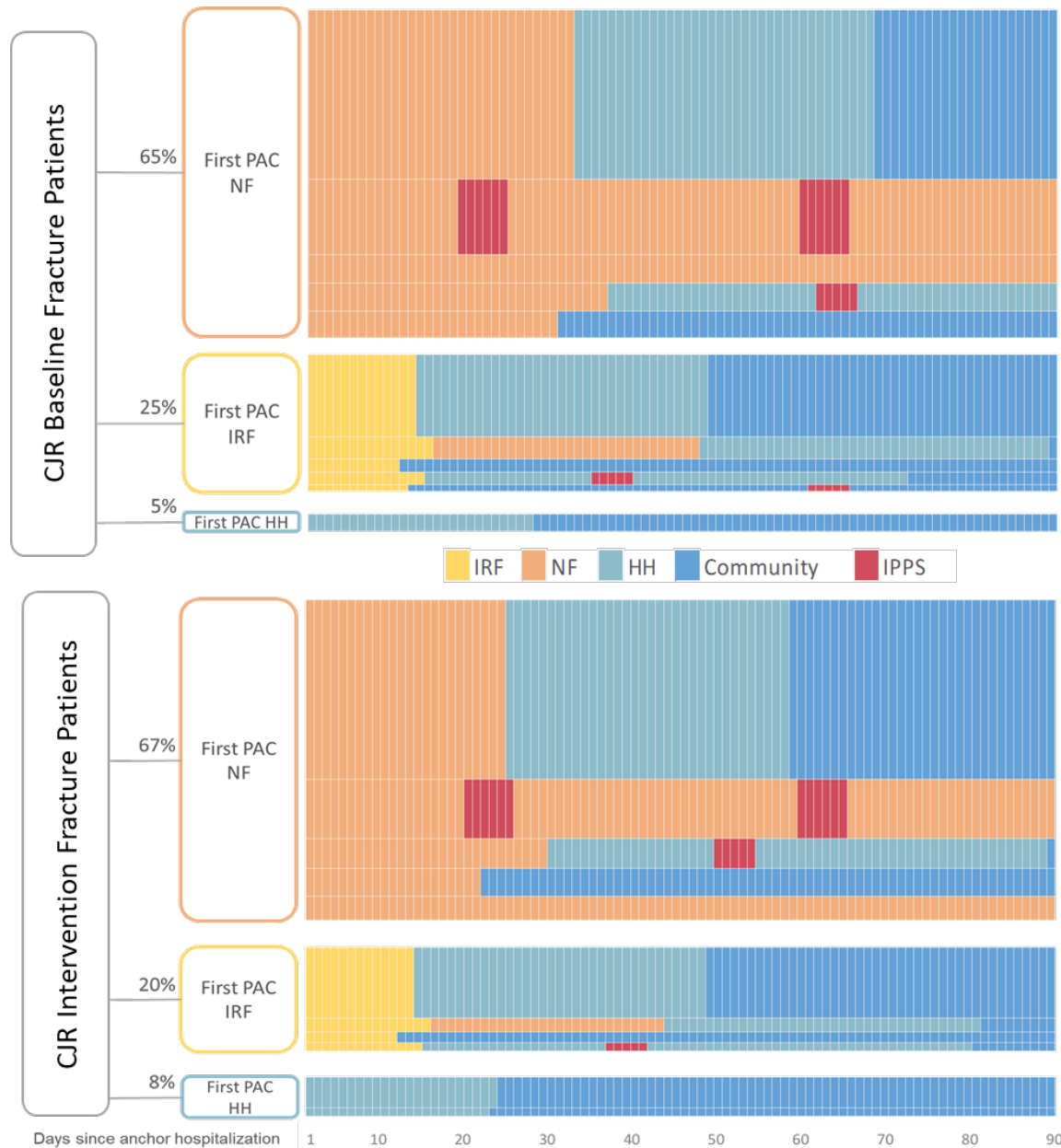
Notes*:* Percentages do not add up to 100 for each first PAC setting as only the most common pathways are depicted. Each row represents a unique care pathway and is graphed with a height representative of the number of fracture patients with that pathway for each period (baseline/intervention). The number of days in each setting was calculated as the median number of days each fracture patient with that pathway stayed in that setting.

Impact of the CJR model on fracture patients after the episode period

Hip fractures are associated with longer term morbidity, including increased risk of a second fracture, dementia, and mortality.[1,2,3] We investigated the impact of the CJR model on the long-term health trajectory of fracture patients by looking at claims-based outcomes beyond the episode period (91-360 days after discharge). These analyses included episodes from PY1 through PY4 to allow sufficient follow-up time in claims data to investigate outcomes one-year post-LEJR.

Generally, there was little evidence that the CJR model impacted payments, utilization, and DME use in the 12 months after discharge (Appendix L, Exhibits L-5, L-6, and L-8). However, the increase in the number of days in the community during the episode (1.5 days; p=0.02) reversed after the episode. Relative to the baseline, CJR fracture patients spent 1.1 fewer days in the community than control patients 271-360 days post-discharge (p=0.08, Exhibit M-3). It is worth noting that CJR and control group hospitals may have had different trends for days in the community during the 271-360 days after their LEJR in the baseline, which means we cannot be certain that we isolated the impact of the CJR model on this outcome. However, the decreasing trend over the post-episode periods, as well as consistency with the readmission outcome (discussed next) lend credence to this finding.

---

[1]   Berry SD, Samelson EJ, Hanna MT, et al. Second hip fracture in older men and women: the Framingham Study. Arch Intern Med 2007; 167(18): 1971-76.

[2]   Huette P, Abou-Arab O, Djebara AE, et al. Risk factors and mortality of patients undergoing hip fracture surgery: a one-year follow-up study. Scientific Reports 2020; 10(1): 1-8.

[3]   Lee HB, Oldham MA, Sieber FE, Oh ES. Impact of delirium after hip fracture surgery on one-year mortality in patients with or without dementia: a case of effect modification. The American Journal of Geriatric Psychiatry 2017; 25(3): 308-315.

Exhibit M-3: PY1-3 impacts on days in community reverses after the episode period
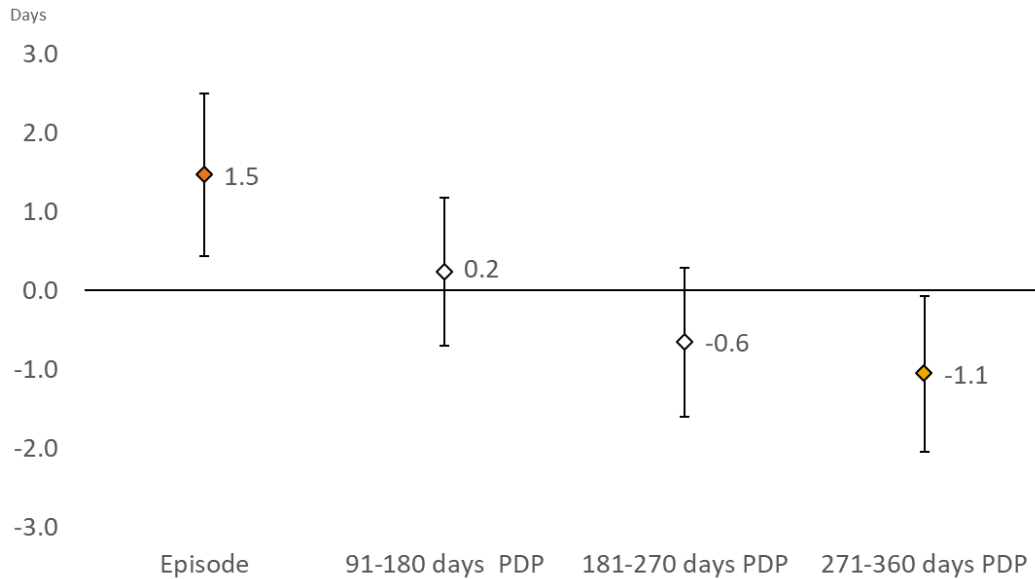
Source: CJR evaluation team analysis of Medicare claims and enrollment data and MDS NF assessment data for hip fracture beneficiaries with LEJR procedures from January 2012 through March 2015 (baseline) and from April 2016 through December 2018 (PY1-3 intervention period). Data includes beneficiaries who died during the given measurement period.

Notes: This analysis was conducted on an earlier intervention period (PY1-3) to avoid using MDS nursing facility assessment data when there was a temporary pause in data reporting due PHE.

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

CJR and control group hospitals may have had different baseline trends for days in community 271-360 days PDP, which means we cannot be certain that we isolated the impact of the CJR model on this outcome. Additional details about baseline trends are included in Appendix C Section III.

There is some indication that CJR fracture patients experienced worse quality of care after their episode. While we found no impact of the CJR model on the unplanned readmission rate for fracture patients during the episode period, the impact of the CJR model on the unplanned readmission rate increased across the post-episode periods, although the impact was only statistically significant in the last post-episode period. During the 271-360 days post-discharge, there was a relative increase of 0.7 pp, or 6.0% of the CJR baseline, in the unplanned readmission rate for CJR fracture patients compared to control patients (p=0.09, Exhibit M-4). This may correspond to the relative decrease in days in the community for the same time period (previously discussed). More hospital readmissions and hospital stays would result in CJR fracture patients spending fewer days in the community.

Exhibit M-4: PY1-4 impacts on unplanned readmission rates for fracture patients appear to be increasing across post-episode periods

Source: CJR evaluation team analysis of Medicare claims and enrollment data for hip fracture beneficiaries with LEJR procedures from January 2012 through March 2015 (baseline) and from April 2016 through December 2019 (PY1-4 intervention period). Data includes beneficiaries who died during the given measurement period.

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

Additionally, CJR fracture patients had a relative increase in incidence of pressure ulcers when measured up to one-year after the LEJR surgery (0.7 pp or 10.9% increase from CJR baseline, p=0.10) (Exhibit M-5). (Due to the low prevalence of pressure ulcers in our data, we could only estimate the relative change in the incidence of pressure ulcers for the episode and one-year post-surgery periods.)

Exhibit M-5: PY1-4 impacts on the incidence of pressure ulcers and delirium for fracture patients

Outcome	Episode period			One-year post surgery		
	CJR intervention risk-adjusted average	DiD (pp)	p-value	CJR intervention risk-adjusted average	DiD (pp)	p-value
Incidence of pressure ulcers	6.1%	0.5	0.21	7.9%	0.7	0.10
Incidence of delirium	3.8%	-0.2	0.42	4.8%	-0.2	0.49

Source: CJR evaluation team analysis of Medicare claims and enrollment data for 90-day episodes initiated in 2012 through 2014 that ended between April 2012 and March 2015 (baseline) and 90-day episodes initiated during or after April 2016 that ended by December 2019 (intervention). Data includes beneficiaries who died during the given measurement period. One-year post surgery outcomes measures from the start of the anchor LEJR hospitalization to 360 days after the anchor discharge.

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

There was generally no impact on mortality, ED use, or delirium after the episode. Between 91-180 days after discharge, CJR fracture patients had a relative reduction in mortality compared to control fracture patients (0.5 pp, p=0.09). However, the impacts on mortality in later post-episode periods were positive, small in magnitude, and statistically insignificant (Exhibit M-6). Because there was not a consistent pattern, we concluded that the CJR model generally did not impact mortality after the episode. Estimates for ED use during post-episode periods were negative, meaning there was a relative reduction in ED use, that was small in magnitude and statistically insignificant (Appendix L, Exhibit L-7). There was no impact on delirium during the one-year post-surgery.

Exhibit M-6: Mortality rates for fracture patients during and beyond the episode period showed no impact of the CJR model, fracture patients, PY1-PY4
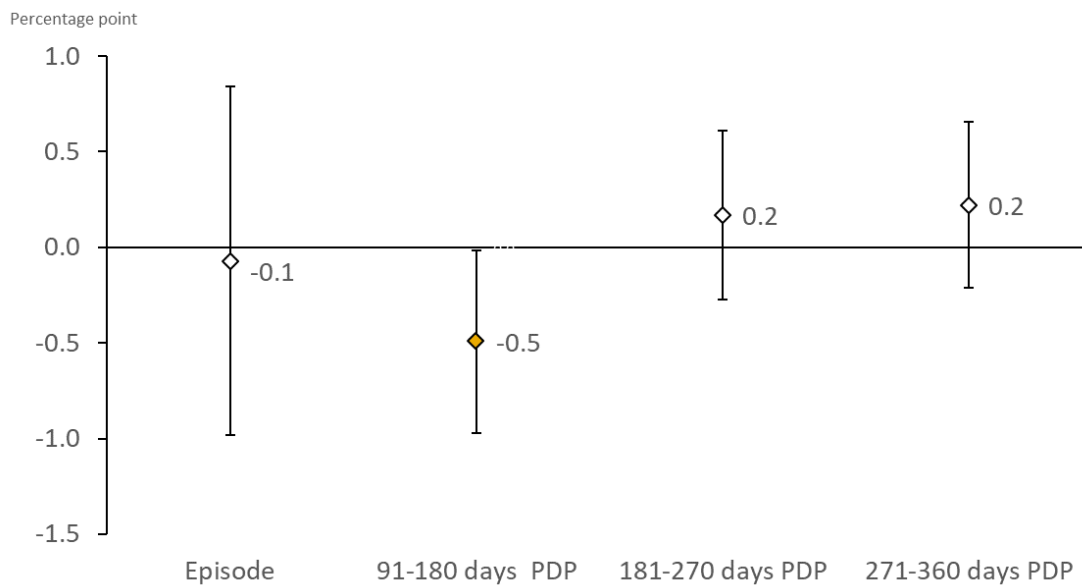
Source: CJR evaluation team analysis of Medicare claims and enrollment data for hip fracture beneficiaries with LEJR procedures from January 2012 through March 2015 (baseline) and from April 2016 through December 2019 (PY1-4 intervention period). Data includes beneficiaries who died during the given measurement period.
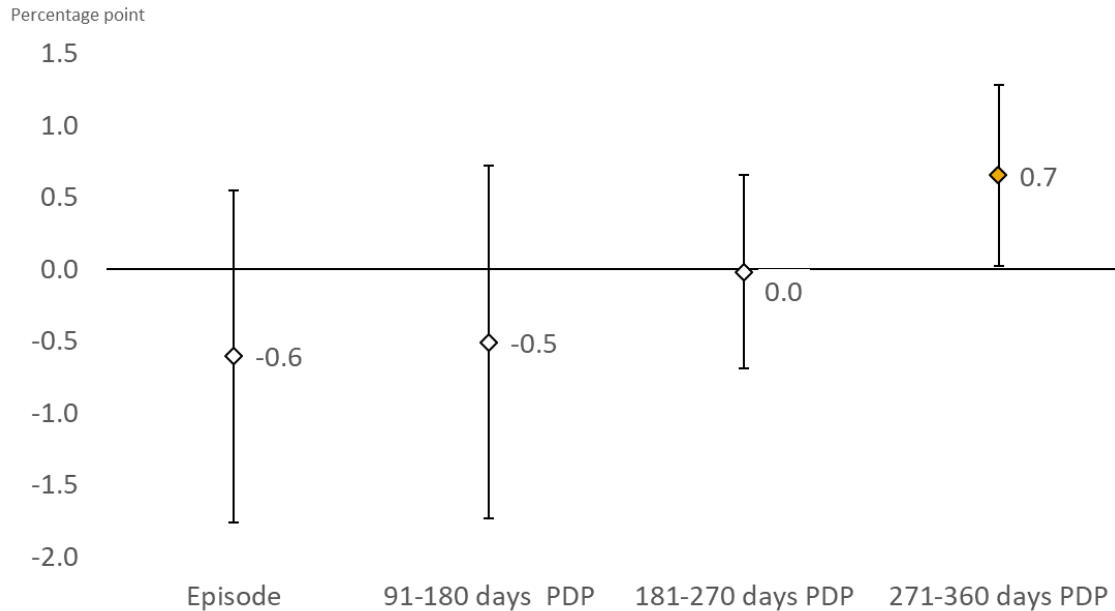
Notes: The estimates in this exhibit are the result of a DiD model. DiD estimates that are significant at the 1%, 5%, or 10% significance level are indicated by red, orange, or yellow shaded circles, respectively. The whiskers represent 90% confidence intervals.

Appendix N: Patient Survey Questions

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 <u>week before</u> 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

Appendix O: Patient Survey Results

Exhibit O-1: Risk-adjusted survey-based results for change in functional status, satisfaction with overall recovery, satisfaction with care management, care transitions, and caregiver help, LEJR patients with hip fractures discharged from mandatory hospitals

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	543	617	-0.68	-0.67	-0.01 (-0.3%)	0.89
	Difficulty walking up or down 12 stairs	-3 to 3	503	583	-0.43	-0.51	0.08 (3.1%)	0.30
	Difficulty rising from sitting	-4 to 4	571	637	-0.23	-0.30	0.07 (1.7%)	0.20
	Difficulty standing	-4 to 4	561	642	-0.16	-0.21	0.05 (1.2%)	0.34
	Use of a mobility aid	-2 to 2	564	632	-0.65	-0.62	-0.03 (-1.4%)	0.34
	Difficulty getting on/off the toilet	-4 to 4	567	639	-0.03	-0.03	-0.01 (-0.2%)	0.93
	Frequency that pain interferes with normal activities	-4 to 4	564	648	-0.36	-0.29	-0.07 (-1.7%)	0.18
	Medication use for pain in the joint you had replaced	-3 to 3	549	621	-0.32	-0.30	-0.02 (-0.6%)	0.62
Satisfaction with overall recovery[b]	Satisfaction with overall recovery since leaving the hospital	0 to 100	618	696	74.5	74.9	-0.5	0.83
Satisfaction with care management[b]	Composite measure of satisfaction with care management	0 to 100	588	666	72.2	75.9	-3.7	0.08
	Health care providers listened to preferences	0 to 100	608	692	67.5	72.3	-4.8	0.05
	Satisfaction with discharge destination	0 to 100	614	696	66.6	72.2	-5.6	0.03
	Satisfaction with care coordination	0 to 100	608	683	71.9	75.1	-3.1	0.16
	Satisfaction with treatment instructions	0 to 100	615	698	73.9	77.8	-3.9	0.11
Experience with care transitions[c]	Discharged from the hospital at the right time	Yes	578	651	84.8	84.5	0.3	0.86
	Received the right amount of post-discharge care	Yes	598	678	76.5	75.6	0.9	0.77
	Had all the medical equipment needed at home	Yes	542	618	88.2	91.1	-2.9	0.08

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]	Yes	552	623	95.9	95.9	0.0	1.00
	Composite measure of caregiver help[d]	0 to 100	538	607	55.7	54.6	1.1	0.62
	Help needed putting on or taking off clothes[d]	0 to 100	546	624	49.4	48.4	1.0	0.64
	Help needed bathing[d]	0 to 100	546	616	50.1	49.8	0.4	0.88
	Help needed using the toilet[d]	0 to 100	544	620	67.2	65.9	1.3	0.61

Source: CJR evaluation team analysis of patient survey data for episodes with discharge in July, August, September, or October 2021.

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.

        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 O-2: Risk-adjusted survey-based results for change in functional status, satisfaction with overall recovery, satisfaction with care management, care transitions, and caregiver help, LEJR patients with hip fractures discharged from mandatory hospitals – 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	543	509	-0.68	-0.67	-0.01 (-0.2%)	0.92
	Difficulty walking up or down 12 stairs	-3 to 3	503	475	-0.44	-0.47	0.03 (1.0%)	0.74
	Difficulty rising from sitting	-4 to 4	571	524	-0.25	-0.28	0.03 (0.7%)	0.60
	Difficulty standing	-4 to 4	561	528	-0.16	-0.19	0.03 (0.7%)	0.62
	Use of a mobility aid	-2 to 2	564	520	-0.67	-0.63	-0.05 (-2.0%)	0.10
	Difficulty getting on/off the toilet	-4 to 4	567	526	-0.04	0.00	-0.04 (-1.0%)	0.57
	Frequency that pain interferes with normal activities	-4 to 4	564	538	-0.34	-0.25	-0.09 (-2.2%)	0.10
	Medication use for pain in the joint you had replaced	-3 to 3	549	512	-0.32	-0.28	-0.04 (-1.0%)	0.49
Satisfaction with overall recovery[b]	Satisfaction with overall recovery since leaving the hospital	0 to 100	618	577	74.3	75.3	-1.0	0.60
Satisfaction with care management[b]	Composite measure of satisfaction with care management	0 to 100	588	550	71.2	76.0	-4.7	0.03
	Health care providers listened to preferences	0 to 100	608	574	67.4	73.0	-5.6	0.04
	Satisfaction with discharge destination	0 to 100	614	578	66.4	72.4	-6.0	0.02
	Satisfaction with care coordination	0 to 100	608	566	71.5	75.2	-3.7	0.11
	Satisfaction with treatment instructions	0 to 100	615	579	73.9	77.4	-3.5	0.15
Experience with care transitions[c]	Discharged from the hospital at the right time	Yes	578	541	84.4	84.5	-0.1	0.97
	Received the right amount of post-discharge care	Yes	598	563	76.1	76.1	0.0	1.00
	Had all the medical equipment needed at home	Yes	542	515	87.9	91.9	-3.9	0.02

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]	Yes	552	521	96.0	96.0	-0.0	0.98
	Composite measure of caregiver help[d]	0 to 100	538	507	55.9	54.5	1.3	0.55
	Help needed putting on or taking off clothes[d]	0 to 100	546	520	50.1	49.3	0.8	0.71
	Help needed bathing[d]	0 to 100	546	514	50.4	51.0	-0.6	0.81
	Help needed using the toilet[d]	0 to 100	544	516	67.1	64.8	2.2	0.36

Source: CJR evaluation team analysis of patient survey data for episodes with discharge in July, August, September, or October 2021.

Notes: The estimates in this exhibit are the result of a cross-sectional regression model, weighted for sampling and nonresponse. Estimates excluded any control LEJR performed by a hospital or physician group practice participating in the Major Joint Replacement of the Lower Extremity bundle of the BPCI Advanced model. Estimates that are significant at the 1%, 5%, or 10% significance level are indicated by red, orange, or yellow shaded cells, respectively.

  BPCI = Bundled Payments for Care Improvement.

  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 O-3: Risk-adjusted survey-based results for satisfaction with care management decomposed by satisfaction level, LEJR patients with hip fractures discharged from mandatory participating hospitals

Response Categories	Health care providers listened to preferences			Satisfaction with discharge destination			Satisfaction with care coordination			Satisfaction with treatment instructions		
	CJR (%)	Control (%)	Difference (pp)	CJR (%)	Control (%)	Difference (pp)	CJR (%)	Control (%)	Difference (pp)	CJR (%)	Control (%)	Difference (pp)
Very satisfied	44.5	47.3	-2.8	47.4	54.2	-6.7	49.8	52.3	-2.5	48.7	54.0	-5.3
Somewhat satisfied	19.8	22.7	-3.0	17.4	15.9	1.5	19.9	22.6	-2.7	24.7	21.6	3.1
Neither satisfied nor dissatisfied	10.5	11.1	-0.6	6.4	5.9	0.5	9.4	5.8	3.5	8.3	11.5	-3.2
Somewhat dissatisfied	11.7	9.7	2.0	11.5	12.6	-1.1	10.2	11.8	-1.5	9.9	7.2	2.7
Very dissatisfied	13.5	9.2	4.3	17.2	11.4	5.8	10.8	7.5	3.2	8.3	5.6	2.7

Source: CJR evaluation team analysis of patient survey data for episodes with discharge in July, August, September, or October 2021.

Notes: The estimates in this exhibit are the result of a cross-sectional multinomial logistic regression model, weighted for sampling and nonresponse.

Each cell indicates the proportion of CJR and control respondents who reported a given level of satisfaction, and the difference in proportions between CJR and control respondents.

LEJR = lower extremity joint replacement, pp = percentage point.

Appendix P: Historically Underserved Populations – Patient Characteristics and Impact Results

Patient Characteristics

Exhibit P-1:   Patient Characteristics for Black and White patients who receive elective LEJRs, baseline

Characteristic Type	Characteristic	CJR Baseline			Control Baseline		
		Black	White	Difference*	Black	White	Difference*
Demographic	Age, 80+	11.2%	23.1%	-11.9	10.3%	20.9%	-10.7
	Sex, Female	73.6%	63.2%	10.5	73.7%	63.6%	10.2
	Race, Black or African American	N.A.	N.A.	N.A.	N.A.	N.A.	N.A.
SDOH	Eligible for Medicaid	39.0%	7.1%	31.9	35.8%	6.7%	29.1
	Disability, no ESRD	45.4%	13.7%	31.7	47.4%	13.9%	33.5
Health status	MS-DRG 469	3.4%	2.8%	0.7	3.1%	2.7%	0.4
	HCC Score	1.47	1.25	0.22	1.45	1.19	0.26
	Obesity	29.6%	16.3%	13.3	31.1%	17.0%	14.2
	Diabetes	43.2%	27.2%	16.0	41.0%	25.7%	15.3
	Hypertension	84.7%	74.2%	10.6	86.1%	74.5%	11.5
	Dementia	2.9%	3.3%	-0.4	3.2%	3.4%	-0.2
	CHF	15.5%	7.1%	8.4	16.9%	7.7%	9.1
Prior Use	IPPS stay	12.6%	11.3%	1.3	12.8%	11.2%	1.6
	HH Use	13.5%	9.9%	3.6	18.1%	9.1%	9.0
	IRF Stay	1.2%	1.1%	0.1	1.5%	1.1%	0.4
	SNF Stay	4.2%	3.7%	0.6	3.3%	3.4%	-0.1
	Any prior care	37.8%	25.2%	12.6	41.2%	25.1%	16.1

Source: CJR evaluation team analysis of Medicare claims and enrollment data for LEJR episodes initiated in 2012 through 2014 that ended between April 2012 and March 2015 (baseline).

Notes: * Differences for characteristics other than HCC score are in percentage points. For HCC score, they are in points

Exhibit P-2:   Patient Characteristics for duals and nonduals who receive elective LEJRs, baseline

Characteristic Type	Characteristic	CJR Baseline			Control Baseline		
		Dual	Nondual	Difference*	Dual	Nondual	Difference*
Demographic	Age, 80+	14.1%	22.8%	-8.7	10.1%	20.8%	-10.7
	Sex, Female	77.2%	62.6%	14.5	77.2%	62.9%	14.2
	Race, Black or African American	20.7%	4.9%	15.8	27.0%	5.6%	21.4
SDOH	Eligible for Medicaid	N.A.	N.A.	N.A.	N.A.	N.A.	N.A.
	Disability, no ESRD	49.4%	11.9%	37.5	60.0%	12.0%	47.9
Health status	MS-DRG 469	4.7%	2.5%	2.2	4.3%	2.6%	1.8
	HCC Score	1.81	1.20	0.61	1.77	1.14	0.63
	Obesity	29.8%	15.8%	14.0	32.5%	16.6%	15.9
	Diabetes	43.5%	27.7%	15.8	39.6%	26.1%	13.4
	Hypertension	80.1%	74.7%	5.4	78.5%	75.1%	3.3
	Dementia	6.5%	2.9%	3.6	5.1%	3.2%	1.9
	CHF	14.6%	6.8%	7.8	17.6%	7.5%	10.1
Prior Use	IPPS stay	17.0%	10.6%	6.5	18.4%	10.5%	7.8
	HH Use	21.5%	9.2%	12.3	22.4%	8.6%	13.8
	IRF Stay	1.6%	1.0%	0.6	1.8%	1.1%	0.8
	SNF Stay	7.2%	3.2%	3.9	6.2%	3.0%	3.2
	Any prior care	46.5%	23.8%	22.7	50.6%	23.8%	26.8

Source: CJR evaluation team analysis of Medicare claims and enrollment data for LEJR episodes initiated in 2012 through 2014 that ended between April 2012 and March 2015 (baseline).

Notes: * Differences for characteristics other than HCC score are in percentage points. For HCC score, they are in points.

Exhibit P-3:   Patient Characteristics for Black duals and White nonduals who receive elective LEJRs, baseline

Characteristic Type	Characteristic	CJR Baseline			Control Baseline		
		Black dual	White nondual	Difference*	Black dual	White nondual	Difference*
Demographic	Age, 80+	7.5%	23.8%	-16.3	6.6%	21.6%	-15.1
	Sex, Female	79.3%	62.2%	17.1	79.0%	62.6%	16.4
	Race, Black or African American	N.A.	N.A.	N.A.	N.A.	N.A.	N.A.
SDOH	Eligible for Medicaid	N.A.	N.A.	N.A.	N.A.	N.A.	N.A.
	Disability, no ESRD	69.4%	10.5%	58.9	72.2%	10.5%	61.7
Health status	MS-DRG 469	4.9%	2.6%	2.3	4.0%	2.6%	1.5
	HCC Score	1.82	1.20	0.62	1.80	1.14	0.66
	Obesity	37.3%	15.3%	22.0	39.0%	15.9%	23.1
	Diabetes	45.3%	26.5%	18.8	43.9%	25.0%	18.9
	Hypertension	84.2%	74.1%	10.1	85.7%	74.5%	11.2
	Dementia	4.1%	3.0%	1.1	3.5%	3.2%	0.3
	CHF	19.9%	6.4%	13.4	22.4%	7.1%	15.3
Prior Use	IPPS stay	16.8%	10.7%	6.1	17.7%	10.6%	7.1
	HH Use	19.6%	9.1%	10.4	28.6%	8.4%	20.2
	IRF Stay	1.6%	1.0%	0.5	2.0%	1.1%	1.0
	SNF Stay	5.8%	3.3%	2.6	4.4%	3.1%	1.3
	Any prior care	52.5%	23.5%	28.9	58.5%	23.4%	35.1

Source: CJR evaluation team analysis of Medicare claims and enrollment data for LEJR episodes initiated in 2012 through 2014 that ended between April 2012 and March 2015 (baseline).

Notes: * Differences for characteristics other than HCC score are in percentage points. For HCC score, they are in points

Payment Outcomes

Exhibit P-4: Total Payments

Population		Historically underserved population				Reference population				Impact of the CJR model on							
		CJR		Control		CJR		Control		Historically underserved population		Reference population		Differential Impact		Baseline gap	% change in baseline gap
Historically under-served	Reference	Baseline (B)	Intervention (I)	B	I	B	I	B	I	Estimate	P-value	Estimate	P-value	Estimate	P-value		
Black	White	$30,664	$27,395	$28,994	$28,105	$28,976	$26,319	$28,631	$27,332	-$2,380	0.00	-$1,358	0.00	-$1,023	0.06	$1,689	-60.6%
Dual	Nondual	$35,492	$33,123	$34,048	$33,509	$28,292	$25,503	$27,895	$26,516	-$1,830	0.00	-$1,410	0.00	-$420	0.30	$7,199	-5.8%
Black dual	White nondual	$33,144	$30,636	$30,882	$30,683	$28,308	$25,593	$27,914	$26,548	-$2,309	0.01	-$1,349	0.00	-$961	0.17	$4,836	-19.9%

Source: CJR evaluation team analysis of Medicare claims and enrollment data for LEJR episodes initiated in 2012 through 2014 that ended between April 2012 and March 2015 (baseline) and LEJR episodes initiated during or after April 2016 that ended by December 2020 (intervention).

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

Utilization Outcomes

Exhibit P-5: First PAC – Inpatient Rehabilitation Facilities (IRF)

Population		Historically underserved population				Reference population				Impact of the CJR model on							
		CJR		Control		CJR		Control		Historically underserved population		Reference population		Differential Impact		Baseline gap	% change in baseline gap
Historically under-served	Reference	Baseline (B)	Intervention (I)	B	I	B	I	B	I	Estimate	P-value	Estimate	P-value	Estimate	P-value		
Black	White	17.1%	6.9%	14.1%	9.5%	13.6%	5.9%	13.0%	9.1%	-5.61	0.00	-3.76	0.01	-1.86	0.15	3.58	-51.9%
Dual	Nondual	15.8%	7.7%	13.5%	9.4%	13.7%	5.8%	13.1%	9.1%	-3.88	0.00	-3.89	0.01	0.01	0.99	2.12	0.4%
Black dual	White nondual	17.7%	7.7%	14.1%	9.0%	13.6%	5.8%	13.0%	9.1%	-4.95	0.02	-3.86	0.01	-1.09	0.54	4.19	-25.9%

Source: CJR evaluation team analysis of Medicare claims and enrollment data for LEJR episodes initiated in 2012 through 2014 that ended between April 2012 and March 2015 (baseline) and LEJR episodes initiated during or after April 2016 that ended by December 2020 (intervention).

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

Exhibit P-6: First PAC – Skilled Nursing Facility (SNF)

Population		Historically underserved population				Reference population				Impact of the CJR model on						Baseline gap	% change in baseline gap
		CJR		Control		CJR		Control		Historically underserved population		Reference population		Differential Impact			
Historically under-served	Reference	Baseline (B)	Intervention (I)	B	I	B	I	B	I	Estimate	P-value	Estimate	P-value	Estimate	P-value		
Black	White	43.9%	32.0%	43.4%	36.2%	40.2%	28.1%	41.5%	31.5%	-4.75	0.13	-2.18	0.17	-2.57	0.28	3.71	-69.4%
Dual	Nondual	53.5%	44.9%	52.1%	46.6%	38.9%	26.3%	40.1%	29.9%	-3.13	0.08	-2.38	0.15	-0.75[a]	0.54	14.57	-5.2%
Black dual	White nondual	44.8%	35.9%	44.0%	38.8%	38.9%	26.4%	40.0%	29.7%	-3.71	0.26	-2.13[b]	0.19	-1.59	0.58	5.95	-26.7%

Source: CJR evaluation team analysis of Medicare claims and enrollment data for LEJR episodes initiated in 2012 through 2014 that ended between April 2012 and March 2015 (baseline) and LEJR episodes initiated during or after April 2016 that ended by December 2020 (intervention).

Notes: Estimates that are significant at the 1%, 5%, or 10% significance level are indicated by red, orange, or yellow shaded cells, respectively. [a] We cannot be certain that we isolated the differential impact of the CJR model on this outcome due to differences in baseline trends. [b] CJR and control group hospitals may have had different trends in the baseline, which means we cannot be certain that we isolated the impact of the CJR model on this outcome.

Exhibit P-7: First PAC – Home Health (HH)

Population		Historically underserved population				Reference population				Impact of the CJR model on						Baseline gap	% change in baseline gap
		CJR		Control		CJR		Control		Historically underserved population		Reference population		Differential Impact			
Historically under-served	Reference	Baseline (B)	Intervention (I)	B	I	B	I	B	I	Estimate	P-value	Estimate	P-value	Estimate	P-value		
Black	White	32.8%	46.3%	31.9%	37.0%	37.5%	50.0%	34.0%	39.7%	8.44	0.03	6.80	0.07	1.64	0.49	-4.69	-34.9%
Dual	Nondual	25.2%	37.2%	27.2%	32.4%	38.6%	51.4%	34.9%	40.6%	6.78	0.00	7.10	0.07	-0.32	0.90	-13.44	2.4%
Black dual	White nondual	31.8%	43.1%	32.1%	37.0%	38.6%	51.3%	35.1%	40.9%	6.34	0.05	6.89[a]	0.07	-0.55	0.88	-6.85	8.0%

Source: CJR evaluation team analysis of Medicare claims and enrollment data for LEJR episodes initiated in 2012 through 2014 that ended between April 2012 and March 2015 (baseline) and LEJR episodes initiated during or after April 2016 that ended by December 2020 (intervention).

Notes: Estimates that are significant at the 1%, 5%, or 10% significance level are indicated by red, orange, or yellow shaded cells, respectively. [a] CJR and control group hospitals may have had different trends in the baseline, which means we cannot be certain that we isolated the impact of the CJR model on this outcome.

LEWINGROUP®

Exhibit P-8: First PAC – Home without HH

Population		Historically underserved population				Reference population				Impact of the CJR model on							
		CJR		Control		CJR		Control		Historically underserved population		Reference population		Differential Impact		Baseline gap	% change in baseline gap
Historically under-served	Reference	Baseline (B)	Intervention (I)	B	I	B	I	B	I	Estimate	P-Value	Estimate	P-value	Estimate	P-value		
Black	White	6.1%	14.8%	10.6%	17.3%	8.7%	16.0%	11.5%	19.7%	1.93	0.39	-0.87	0.73	2.80	0.05	-2.60	-107.7%
Dual	Nondual	5.5%	10.2%	7.2%	11.6%	8.8%	16.4%	11.9%	20.4%	0.23	0.85	-0.83	0.75	1.06	0.58	-3.26	-32.6%
Black dual	White nondual	5.6%	13.3%	9.8%	15.2%	8.9%	16.5%	11.8%	20.3%	2.32	0.25	-0.90	0.73	3.22	0.16	-3.28	-98.2%

Source: CJR evaluation team analysis of Medicare claims and enrollment data for LEJR episodes initiated in 2012 through 2014 that ended between April 2012 and March 2015 (baseline) and LEJR episodes initiated during or after April 2016 that ended by December 2020 (intervention).

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

Exhibit P-9: IRF Days

Population		Historically underserved population				Reference population				Impact of the CJR model on							
		CJR		Control		CJR		Control		Historically underserved population		Reference population		Differential Impact		Baseline gap	% change in baseline gap
Historically under-served	Reference	Baseline (B)	Intervention (I)	B	I	B	I	B	I	Estimate	P-value	Estimate	P-value	Estimate	P-value		
Black	White	11.4	11.7	11.4	11.7	11.8	12.0	11.6	12.0	-0.05	0.88	-0.14	0.48	0.09	0.76	-0.38	-24.1%
Dual	Nondual	12.1	12.1	12.0	12.2	11.7	12.0	11.6	11.9	-0.24	0.39	-0.12	0.53	-0.12	0.63	0.32	-35.5%
Black dual	White nondual	11.4	11.6	11.4	11.2	11.8	12.0	11.6	11.9	0.31	0.51	-0.14	0.47	0.45	0.37	-0.37	-125.0%

Source: CJR evaluation team analysis of Medicare claims and enrollment data for LEJR episodes initiated in 2012 through 2014 that ended between April 2012 and March 2015 (baseline) and LEJR episodes initiated during or after April 2016 that ended by December 2020 (intervention).

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

LEWINGROUP®

Exhibit P-10: SNF Days

Population		Historically underserved population				Reference population				Impact of the CJR model on				Differential Impact		Baseline gap	% change in baseline gap
		CJR		Control		CJR		Control		Historically underserved population		Reference population					
Historically under-served	Reference	Baseline (B)	Intervention (I)	B	I	B	I	B	I	Estimate	P-value	Estimate	P-value	Estimate	P-value	Baseline gap	% change in baseline gap
Black	White	28.1	23.4	26.1	24.8	27.1	22.3	27.4	24.9	-3.37	0.00	-2.36	0.00	-1.01	0.28	0.94	-107.2%
Dual	Nondual	35.2	31.2	34.0	32.2	25.6	20.6	25.8	23.4	-2.18	0.00	-2.46	0.00	0.29	0.61	9.60	3.0%
Black dual	White nondual	32.0	28.5	29.1	28.6	25.6	20.7	25.9	23.4	-3.02	0.03	-2.40	0.00	-0.62	0.59	6.35	-9.8%

Source: CJR evaluation team analysis of Medicare claims and enrollment data for LEJR episodes initiated in 2012 through 2014 that ended between April 2012 and March 2015 (baseline) and LEJR episodes initiated during or after April 2016 that ended by December 2020 (intervention).

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

Exhibit P-11: HH Visits

Population		Historically underserved population				Reference population				Impact of the CJR model on				Differential Impact		Baseline gap	% change in baseline gap
		CJR		Control		CJR		Control		Historically underserved population		Reference population					
Historically under-served	Reference	Baseline (B)	Intervention (I)	B	I	B	I	B	I	Estimate	P-value	Estimate	P-value	Estimate	P-value	Baseline gap	% change in baseline gap
Black	White	18.1	16.0	17.9	17.2	16.8	15.4	16.6	16.2	-1.34[a]	0.01	-1.03[a]	0.02	-0.31	0.41	1.30	-23.9%
Dual	Nondual	19.3	18.0	18.9	18.4	16.7	15.2	16.5	16.0	-0.86	0.19	-1.06[a]	0.03	0.20[b]	0.80	2.67	7.6%
Black dual	White nondual	18.9	16.9	18.5	18.1	16.6	15.2	16.3	15.9	-1.58[a]	0.02	-1.04[a]	0.03	-0.54[b]	0.49	2.33	-23.1%

Source: CJR evaluation team analysis of Medicare claims and enrollment data for LEJR episodes initiated in 2012 through 2014 that ended between April 2012 and March 2015 (baseline) and LEJR episodes initiated during or after April 2016 that ended by December 2020 (intervention).

Notes: Estimates that are significant at the 1%, 5%, or 10% significance level are indicated by red, orange, or yellow shaded cells, respectively. [a] CJR and control group hospitals may have had different trends in the baseline, which means we cannot be certain that we isolated the impact of the CJR model on this outcome. [b] We cannot be certain that we isolated the differential impact of the CJR model on this outcome due to differences in baseline trends.

Quality Outcomes

Exhibit P-12: Mortality

Population		Historically underserved population				Reference population				Impact of the CJR model on						Baseline gap	% change in baseline gap
		CJR		Control		CJR		Control		Historically underserved population		Reference population		Differential Impact			
Historically under-served	Reference	Baseline (B)	Intervention (I)	B	I	B	I	B	I	Estimate	P-value	Estimate	P-value	Estimate	P-value		
Black	White	1.7%	1.5%	1.7%	1.9%	2.9%	2.7%	3.0%	2.7%	-0.33	0.07	0.09	0.38	-0.42	0.03	-1.21	34.9%
Dual	Nondual	4.0%	3.8%	4.6%	4.3%	2.5%	2.4%	2.6%	2.4%	0.18[a]	0.41	0.03	0.75	0.15[b]	0.50	1.41	10.5%
Black dual	White nondual	2.1%	1.8%	2.1%	2.2%	2.6%	2.5%	2.7%	2.5%	-0.44	0.26	0.05	0.60	-0.49	0.23	-0.55	89.2%

Source: CJR evaluation team analysis of Medicare claims and enrollment data for LEJR episodes initiated in 2012 through 2014 that ended between April 2012 and March 2015 (baseline) and LEJR episodes initiated during or after April 2016 that ended by December 2020 (intervention).

Notes: Estimates that are significant at the 1%, 5%, or 10% significance level are indicated by red, orange, or yellow shaded cells, respectively. [a] CJR and control group hospitals may have had different trends in the baseline, which means we cannot be certain that we isolated the impact of the CJR model on this outcome. [b] We cannot be certain that we isolated the differential impact of the CJR model on this outcome due to differences in baseline trends.

Exhibit P-13: Emergency Department (ED) Use

Population		Historically underserved population				Reference population				Impact of the CJR model on						Baseline gap	% change in baseline gap
		CJR		Control		CJR		Control		Historically underserved population		Reference population		Differential Impact			
Historically under-served	Reference	Baseline (B)	Intervention (I)	B	I	B	I	B	I	Estimate	P-value	Estimate	P-value	Estimate	P-value		
Black	White	17.7%	18.9%	17.3%	18.4%	12.8%	13.7%	12.3%	13.2%	0.01[a]	0.99	0.05	0.88	-0.04[b]	0.97	4.96	-0.7%
Dual	Nondual	18.9%	19.9%	19.6%	20.7%	12.1%	13.1%	11.5%	12.4%	-0.07[a]	0.92	0.12	0.66	-0.19	0.75	6.81	-2.8%
Black dual	White nondual	23.6%	24.3%	23.7%	25.4%	12.0%	13.1%	11.6%	12.5%	-1.01[a]	0.50	0.12	0.67	-1.13	0.43	11.55	-9.8%

Source: CJR evaluation team analysis of Medicare claims and enrollment data for LEJR episodes initiated in 2012 through 2014 that ended between April 2012 and March 2015 (baseline) and LEJR episodes initiated during or after April 2016 that ended by December 2020 (intervention).

Notes: Estimates that are significant at the 1%, 5%, or 10% significance level are indicated by red, orange, or yellow shaded cells, respectively. [a] CJR and control group hospitals may have had different trends in the baseline, which means we cannot be certain that we isolated the impact of the CJR model on this outcome. [b] We cannot be certain that we isolated the differential impact of the CJR model on this outcome due to differences in baseline trends.

Exhibit P-14: Readmissions

Population		Historically underserved population				Reference population				Impact of the CJR model on						Baseline gap	% change in baseline gap
		CJR		Control		CJR		Control		Historically underserved population		Reference population		Differential Impact			
Historically under-served	Reference	Baseline (B)	Intervention (I)	B	I	B	I	B	I	Estimate	P-value	Estimate	P-value	Estimate	P-value	Baseline gap	% change in baseline gap
Black	White	10.5%	9.7%	10.1%	9.7%	9.1%	8.9%	8.8%	8.8%	-0.38	0.56	-0.19	0.38	-0.19	0.75	1.41	-13.4
Dual	Nondual	13.5%	13.1%	12.8%	12.9%	8.5%	8.2%	8.2%	8.2%	-0.41	0.44	-0.22	0.30	-0.19	0.71	5.02	-3.7%
Black dual	White nondual	13.7%	13.0%	12.4%	11.7%	8.6%	8.3%	8.3%	8.3%	-0.03	0.98	-0.16	0.43	0.14	0.89	5.18	2.6%

Source: CJR evaluation team analysis of Medicare claims and enrollment data for LEJR episodes initiated in 2012 through 2014 that ended between April 2012 and March 2015 (baseline) and LEJR episodes initiated during or after April 2016 that ended by December 2020 (intervention).

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

LEJR Access Outcomes

Exhibit P-15: Elective LEJR rates per 100,000 beneficiaries for Black beneficiaries, duals, Black duals and their reference populations (White beneficiaries, nonduals, White nonduals, respectively)

Population		Historically underserved population				Reference population				Impact of the CJR model on						Baseline gap	% change in baseline gap
		CJR		Control		CJR		Control		Historically underserved population		Reference population		Differential Impact			
Historically under-served	Reference	Baseline (B)	Intervention (I)	B	I	B	I	B	I	Estimate	P-value	Estimate	P-value	Estimate	P-value	Baseline gap	% change in baseline gap
Black	White	709	641	712	679	1,222	1,175	1,274	1,211	-35	0.31	16	0.69	-51	0.06	-513	9.9%
Dual	Nondual	547	485	610	550	1,244	1,202	1,285	1,229	-1	0.95	14	0.73	-15	0.53	-697	2.2%
Black dual	White nondual	530	448	561	524	1,275	1,235	1,330	1,274	-45	0.13	17	0.67	-62	0.06	-745	8.4%

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

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

Exhibit P-16: Elective LEJR rates per 100,000 beneficiaries for Black duals and Black nonduals

Population		Historically underserved population				Reference population				Impact of the CJR model on						Baseline gap	% change in baseline gap
		CJR		Control		CJR		Control		Historically underserved population		Reference population		Differential Impact			
Historically under-served	Reference	Baseline (B)	Intervention (I)	B	I	B	I	B	I	Estimate	P-value	Estimate	P-value	Estimate	P-value	Baseline gap	% change in baseline gap
Black dual	**Black nondual**	538	439	585	521	848	775	882	823	-36	0.21	-15	0.69	-21	0.50	-311	6.7%

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

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

Patient Mix Outcomes

We examined changes in patient characteristics (age, sex, race, Medicaid eligibility, disability status, health status, and prior health care use) for each of the three underserved populations. We used the DDD approach described in Section II.1.b except that these regressions are not risk adjusted. We restricted the sample to elective LEJRs because they are planned procedures whereas fracture LEJRs are emergent procedures making selection less likely.

The analyses of patient characteristics are consistent with the analyses of the impacts on LEJR rates in Section II.7.e: The share of CJR Black beneficiaries who were dual declined relative to their counterparts in the control group and differentially declined compared to CJR White beneficiaries. The share of CJR duals who were Black declined relative to their counterparts in the control group and differentially declined compared to nonduals.

CJR Black beneficiaries. CJR Black beneficiaries were relatively less likely to be duals (Exhibit P-18). The share of CJR Black beneficiaries who were duals declined by 4.2 percentage points (p<0.01) from the baseline to the intervention period, relative to control Black beneficiaries. There was no statistically significant change in the proportion of CJR White beneficiaries who were duals, compared to control White beneficiaries. The share of CJR beneficiaries who were duals decreased by 3.9 percentage points (p<0.1) more for Black beneficiaries than for White beneficiaries.

However, the decrease in the dual share of CJR Black beneficiaries was in part due to underlying changes in the demographics of the FFS beneficiary population. The dual share of Black FFS beneficiaries decreased in both CJR and control MSAs, but the decrease was greater in CJR MSAs and accounted for about 40% of the decrease in the dual share of CJR Black beneficiaries relative to control

Black beneficiaries and about 46% of the differential decrease in the dual share of CJR Black beneficiaries compared to CJR White beneficiaries.

CJR Black beneficiaries were relatively more likely to be at least 80 years old.[1] However, this finding may be related to the decrease in the dual share since, for this population, dual status and age are strongly negatively correlated.[2]

CJR Black beneficiaries were relatively less likely to have had a prior SNF stay.[3] There were, however, non-statistically significant increases in all the other prior health care use measures (IPPS stay, HH Use, IRF stay, and Any prior care).

None of the other patient characteristics had statistically significant changes relative to control Black beneficiaries or statistically significant differential changes compared to White beneficiaries.

---

[1] The share of CJR Black beneficiaries who were at least 80 years old increased by 1.5 percentage points (p<0.10) relative to control Black beneficiaries. The share of CJR White beneficiaries who were at least 80 years old decreased by 1.0 percentage point (p<0.10) relative to control White beneficiaries. Thus, the share of CJR beneficiaries who were at least 80 years old increased by 2.5 percentage points (p<0.01) more for Black beneficiaries than for White beneficiaries.

[2] For CJR Black beneficiaries, the correlation between dual eligibility status and age is -0.3.

[3] The share of CJR Black beneficiaries who had a prior SNF stay decreased by 0.9 percentage points (p<0.05) relative to control Black beneficiaries. The proportion of CJR White beneficiaries who had a prior SNF stay decreased by 0.3 percentage points (p<0.01) relative to control White beneficiaries. The differential change for CJR Black beneficiaries and CJR White beneficiaries was not statistically significant.

LEWIN GROUP®

Exhibit P-17: The dual share of CJR Black beneficiaries declined relative to control Black beneficiaries and differentially declined compared to CJR White beneficiaries (Elective LEJRs)
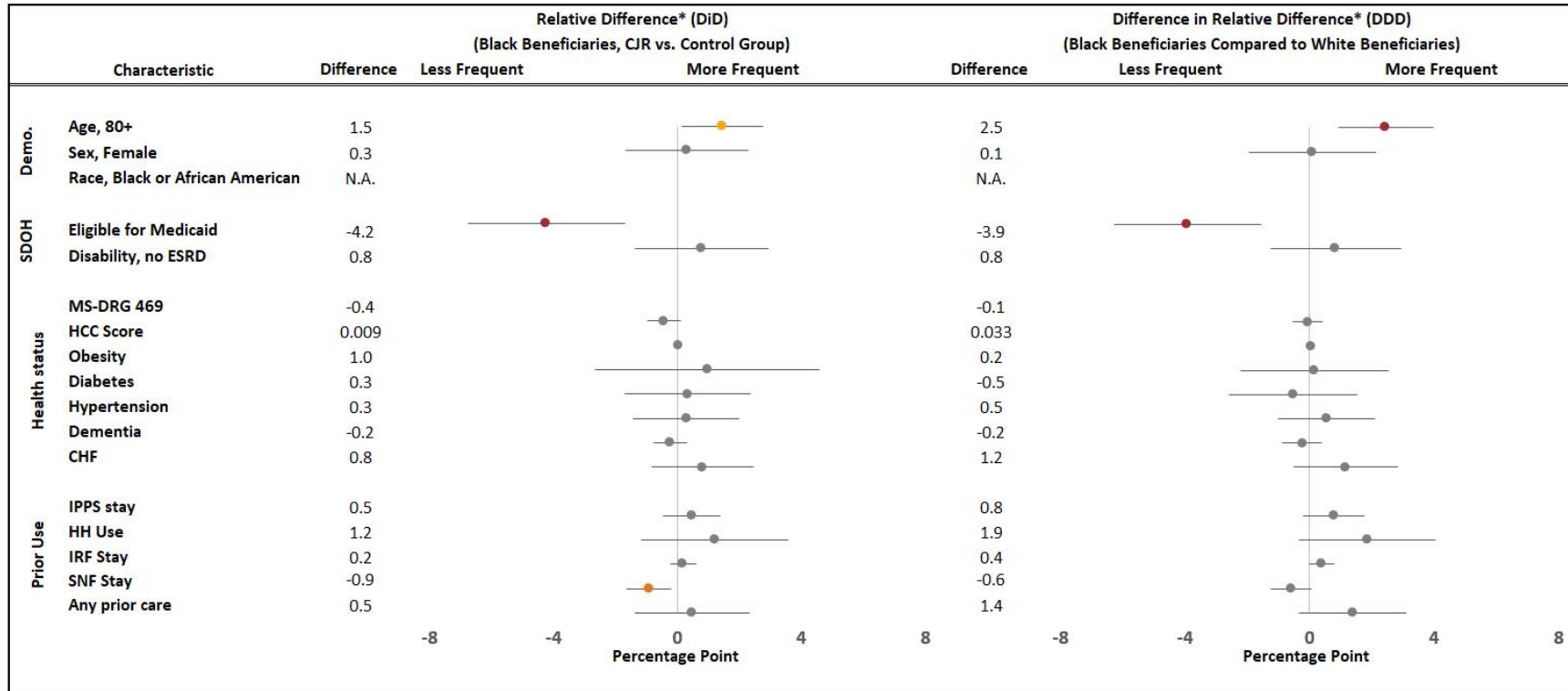
	Relative Difference* (DiD) (Black Beneficiaries, CJR vs. Control Group)		Difference in Relative Difference* (DDD) (Black Beneficiaries Compared to White Beneficiaries)
Characteristic	Difference		Difference
Demo.			
Age, 80+	1.5		2.5
Sex, Female	0.3		0.1
Race, Black or African American	N.A.		N.A.
SDOH			
Eligible for Medicaid	-4.2		-3.9
Disability, no ESRD	0.8		0.8
Health status			
MS-DRG 469	-0.4		-0.1
HCC Score	0.009		0.033
Obesity	1.0		0.2
Diabetes	0.3		-0.5
Hypertension	0.3		0.5
Dementia	-0.2		-0.2
CHF	0.8		1.2
Prior Use			
IPPS stay	0.5		0.8
HH Use	1.2		1.9
IRF Stay	0.2		0.4
SNF Stay	-0.9		-0.6
Any prior care	0.5		1.4

Source: CJR evaluation team analysis of Medicare claims and enrollment data for LEJR episodes initiated in 2012 through 2014 that ended between April 2012 and March 2015 (baseline) and LEJR episodes initiated during or after April 2016 that ended by December 2020 (intervention).

Notes: Differences for all characteristics, except for the HCC score, are in percentage points. For the HCC score, they are in points. Estimates that are significant at the 1%, 5%, or 10% significance level are indicated by red, orange, or yellow shaded points, respectively.

CJR duals. CJR duals were relatively less likely to be Black (Exhibit P-19). The Black share of CJR duals declined by 2.0 percentage points (p<0.10), from the baseline to the intervention period, relative to control duals. There was no statistically significant change in the Black share of CJR nonduals, relative to control nonduals. The Black share of duals decreased by 1.9 percentage points (p<0.10) more than the Black share of nonduals.

CJR duals were relatively less likely to be classified as having major complications and comorbidities (MS-DRG-469).[4] The remaining health status characteristics show no statistically significant changes for CJR duals relative to control duals, although, compared to non-duals, there were statistically significant differential reductions in the proportion of duals with hypertension and dementia.[5]

CJR duals were relatively less likely to have had a prior SNF stay.[6] However, there were non-statistically significant increases in other prior health care use measures (IPPS stay, HH Use, and Any prior care).

None of the other patient characteristics had statistically significant changes relative to control duals or statistically significant differential changes compared to nonduals.
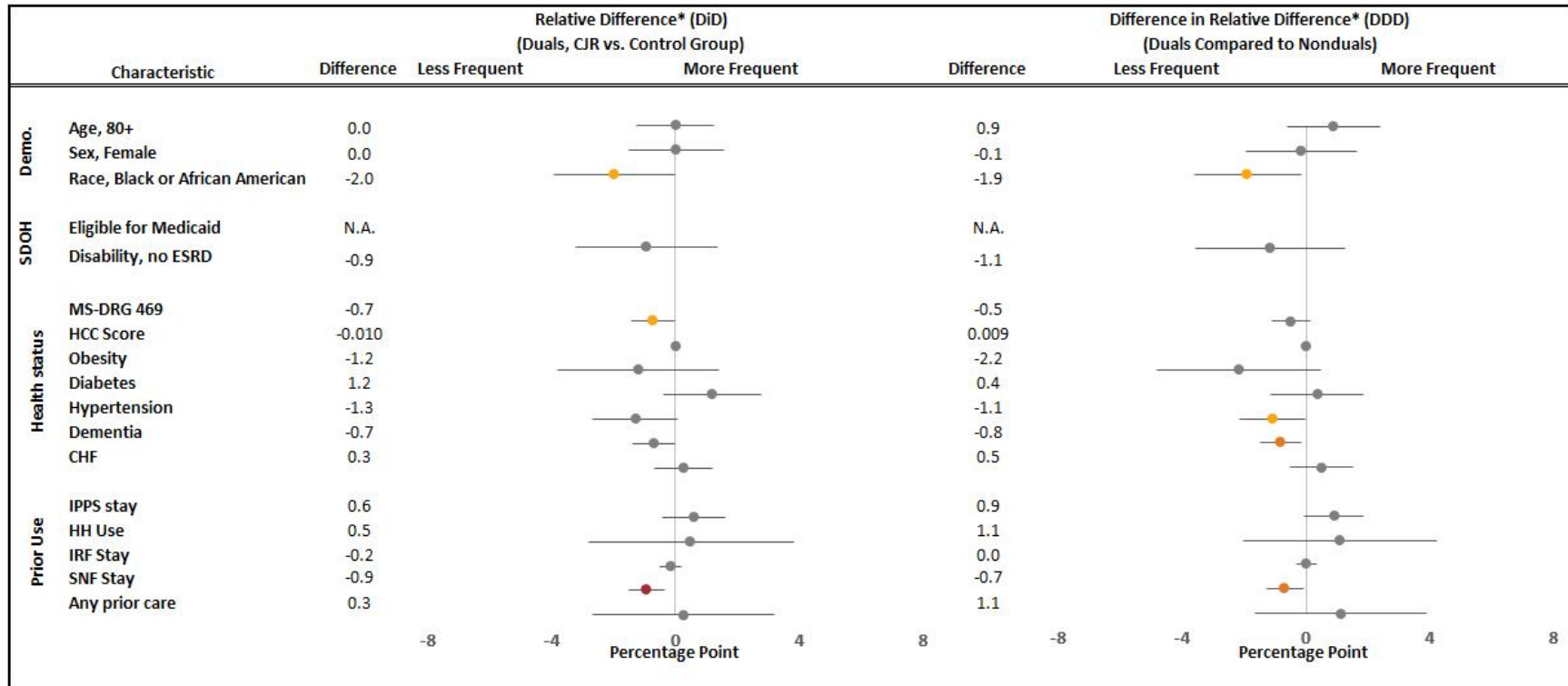
---

[4]   The share of CJR duals classified as having major complications and comorbidities (MS-DRG-469) declined by 0.7 percentage points (p<0.10), relative to control duals. There was no statistically significant change in the share of CJR nonduals classified as MS-DRG-469, relative to control nonduals. The differential change for duals and nonduals was not statistically significant.

[5]   The share of CJR duals with hypertension decreased by 1.1 percentage points more (p<0.10) than the share of CJR nonduals with hypertension, and the share of CJR duals with dementia decreased by 0.8 percentage points more (p<0.05) than the share of CJR nonduals with dementia.

[6]   The share of CJR duals who had a prior SNF stay decreased by 0.9 percentage points (p<0.01), relative to control duals. The share of CJR nonduals who had a prior SNF stay decreased by 0.3 percentage points (p<0.05) relative to control nonduals. The combined effect was to reduce the share of duals who had a prior SNF stay by 0.7 percentage points (p<0.10) more than the share of nonduals who had a prior SNF stay.

LEWINGROUP®

Exhibit P-18: The Black share of CJR duals declined relative to control duals and differentially declined compared to CJR nonduals (Elective LEJRs)

Characteristic	Relative Difference* (DiD) (Duals, CJR vs. Control Group) Difference	Difference in Relative Difference* (DDD) (Duals Compared to Nonduals) Difference
Demo.		
Age, 80+	0.0	0.9
Sex, Female	0.0	-0.1
Race, Black or African American	-2.0	-1.9
SDOH		
Eligible for Medicaid	N.A.	N.A.
Disability, no ESRD	-0.9	-1.1
Health status		
MS-DRG 469	-0.7	-0.5
HCC Score	-0.010	0.009
Obesity	-1.2	-2.2
Diabetes	1.2	0.4
Hypertension	-1.3	-1.1
Dementia	-0.7	-0.8
CHF	0.3	0.5
Prior Use		
IPPS stay	0.6	0.9
HH Use	0.5	1.1
IRF Stay	-0.2	0.0
SNF Stay	-0.9	-0.7
Any prior care	0.3	1.1

Source: CJR evaluation team analysis of Medicare claims and enrollment data for LEJR episodes initiated in 2012 through 2014 that ended between April 2012 and March 2015 (baseline) and LEJR episodes initiated during or after April 2016 that ended by December 2020 (intervention).

Notes: Differences for all characteristics, except for the HCC score, are in percentage points. For the HCC score, they are in points. Estimates that are significant at the 1%, 5%, or 10% significance level are indicated by red, orange, or yellow shaded points, respectively.

CJR Black duals. There were no systematic changes in patient characteristics for CJR Black duals relative to control Black duals. There were also no systematic differential changes in patient characteristics for CJR Black duals compared to White nonduals (Exhibit P-20).

CJR Black duals were relatively less likely to be classified as having major complications and comorbidities (MS-DRG-469).[7] The remaining health status characteristics show no statistically significant changes for CJR Black duals relative to control Black duals. However, compared to White nonduals, there was a statistically significant differential reduction in the share of Black duals with dementia.[8]

CJR Black duals were relatively less likely to have had a prior SNF stay.[9] However, there were non-statistically significant increases in all of the other prior health care use measures (IPPS stay, HH Use, IRF stay, and Any prior care). In addition, compared to White nonduals, there were statistically significant differential increases in the share of Black duals who had prior HH use and IRF stays.[10]

---

[7]   The share of CJR Black duals classified as MS-DRG-469 declined by 1.2 percentage points (p<0.05), relative to control Black duals. There was no statistically significant change in the share of CJR White nonduals classified as MS-DRG-469, relative to control White nonduals. The differential change for Black duals and White nonduals was not statistically significant.

[8]   The share of CJR Black duals with dementia decreased by 0.9 percentage points more (p<0.10) than the share for white nonduals.

[9]   The share of CJR Black duals who had a prior SNF stay decreased by 1.6 percentage points (p<0.10), relative to control Black duals. The share of CJR White nonduals who had a prior SNF stay decreased by 0.3 percentage points (p<0.05), relative to control White nonduals. The differential change for Black duals and White nonduals was not statistically significant.

[10]   The share of Black duals who had prior HH use increased by 3.7 percentage points more (p<0.10) than the share for White nonduals, and the share of Black duals who had a prior IRF stay increased by 0.6 percentage points more (p<0.10) than the share for White nonduals.

Exhibit P-19: There are no systematic changes in patient characteristics for CJR Black duals relative to control Black duals and no systematic differential changes compared to White nonduals (Elective LEJRs)
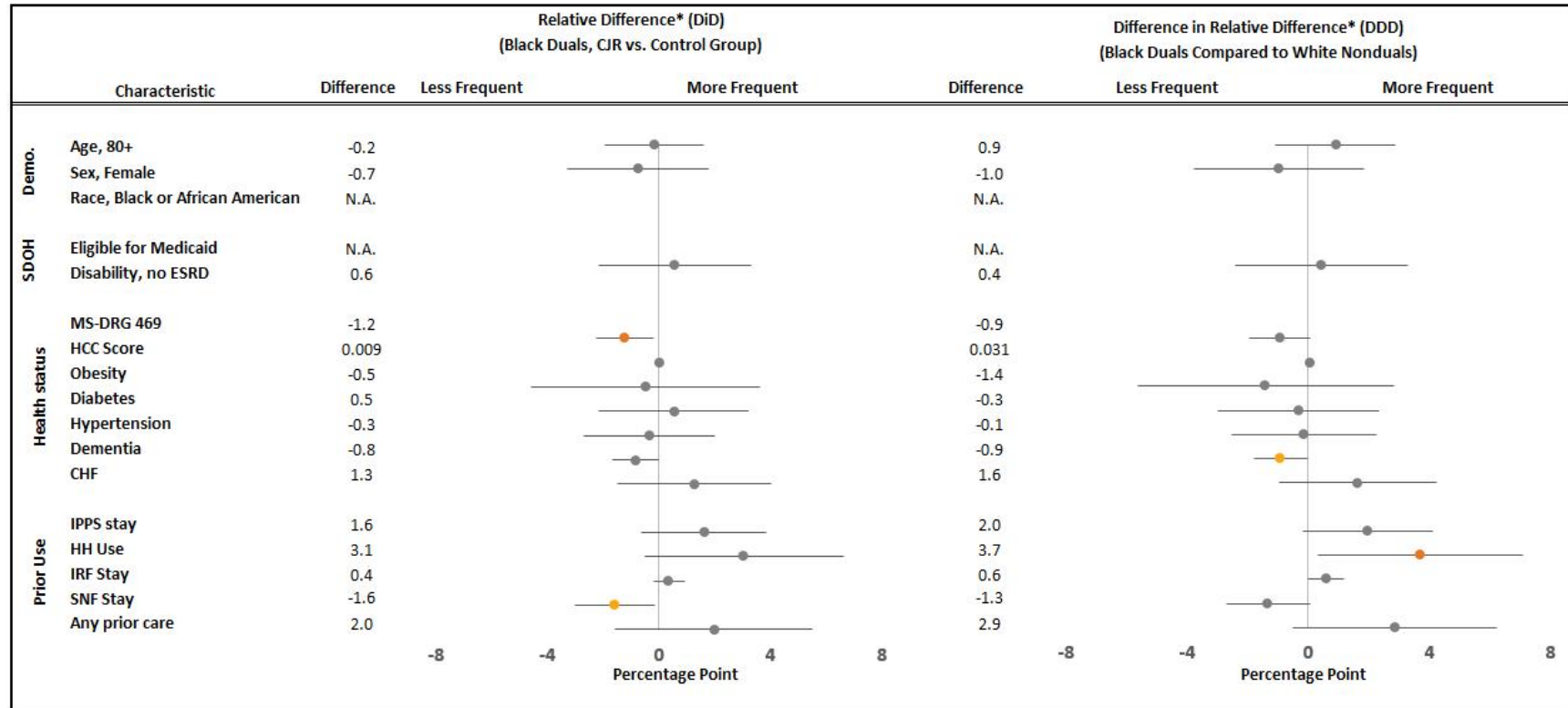
Source: CJR evaluation team analysis of Medicare claims and enrollment data for LEJR episodes initiated in 2012 through 2014 that ended between April 2012 and March 2015 (baseline) and LEJR episodes initiated during or after April 2016 that ended by December 2020 (intervention).

Notes: Differences for all characteristics, except for the HCC score, are in percentage points. For the HCC score, they are in points. Estimates that are significant at the 1%, 5%, or 10% significance level are indicated by red, orange, or yellow shaded points, respectively.

Results of parallel trends tests – Differential Impact Trends

Exhibit P-20: Total Payments

		Diff Impact Trends Episode Payments		
Historically Underserved	**Reference**	**Historically Underserved Population**	**Reference Population**	**DDD**
Black	**White**	p=0.15	p=0.94	p=0.16
Dual	**Nondual**	p=0.29	p=0.46	p=0.45
Black dual	**White nondual**	p=0.19	p=0.57	p=0.28

Source: CJR evaluation team analysis of Medicare claims and enrollment data for LEJR episodes initiated in 2012 through 2014 that ended between April 2012 and March 2015 (baseline).

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

Exhibit P-21: Utilization Outcomes

Utilization Outcome	Historically Underserved	Reference	Diff Impact Trends		
			Historically Underserved Population	Reference Population	DDD
First PAC - IRF	Black	White	p=0.67	p=0.63	p=0.90
	Dual	Nondual	p=0.92	p=0.40	p=0.63
	Black dual	White nondual	p=0.55	p=0.38	p=0.58
First PAC - SNF	Black	White	p=0.72	p=0.27	p=0.42
	Dual	Nondual	p=0.38	p=0.11	p=0.10
	Black dual	White nondual	p=0.58	p=0.08	p=0.43
First PAC - HH	Black	White	p=0.22	p=0.12	p=0.48
	Dual	Nondual	p=0.81	p=0.11	p=0.36
	Black dual	White nondual	p=0.66	p=0.08	p=0.81
First PAC – Home without HH	Black	White	p=0.56	p=0.60	p=0.76
	Dual	Nondual	p=0.66	p=0.68	p=0.50
	Black dual	White nondual	p=0.35	p=0.54	p=0.58
IRF Days	Black	White	p=0.86	p=0.88	p=0.69
	Dual	Non-dual	p=0.83	p=0.66	p=0.39
	Black dual	White nondual	p=0.69	p=0.55	p=0.83
SNF Days	Black	White	p=0.42	p=0.90	p=0.52
	Dual	Nondual	p=0.76	p=0.54	p=0.77
	Black dual	White nondual	p=0.51	p=0.55	p=0.60
HH Visits	Black	White	p=0.06	p=0.00	p=0.77
	Dual	Nondual	p=0.67	p=0.00	p=0.10
	Black dual	White nondual	p=0.03	p=0.00	p=0.04

Source:　CJR evaluation team analysis of Medicare claims and enrollment data for LEJR episodes initiated in 2012 through 2014 that ended between April 2012 and March 2015 (baseline).

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

Exhibit P-22: Quality Outcomes

Quality Outcome	Historically Underserved	Reference	Diff Impact Trends		
			Historically Underserved Population	Reference Population	DDD
Mortality	Black	White	p=0.56	p=0.63	p=0.48
	Dual	Non-dual	p=0.02	p=0.68	p=0.01
	Black dual	White nondual	p=0.36	p=0.76	p=0.27
ED Use	Black	White	p=0.01	p=0.99	p=0.01
	Dual	Nondual	p=0.07	p=0.59	p=0.12
	Black dual	White nondual	p=0.07	p=0.96	p=0.14
Readmissions	Black	White	p=0.52	p=0.43	p=0.53
	Dual	Nondual	p=0.76	p=0.54	p=0.91
	Black dual	White nondual	p=0.51	p=0.33	p=0.59

Source: CJR evaluation team analysis of Medicare claims and enrollment data for LEJR episodes initiated in 2012 through 2014 that ended between April 2012 and March 2015 (baseline).

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

Results of Parallel Trends Tests – LEJR Access Outcomes

Exhibit P-23: LEJR Rate Trends

Historically Underserved	Reference	DDD Pre-trend Tests		
		Historically Underserved Population	Reference Population	DDD
Black	**White**	p=0.82	p=0.59	p=0.99
Dual	**Nondual**	p=0.34	p=0.62	p=0.14
Black dual	**White nondual**	p=0.94	p=0.71	p=0.77

Source: CJR evaluation team analysis of Medicare claims and enrollment data for LEJR episodes initiated in 2012 through 2014 that ended between April 2012 and March 2015 (baseline).

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