



CMS Comprehensive Care for Joint Replacement Model: Performance Years 6–8 Evaluation Report

Final Evaluation Report Appendices

September 2026

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

Submitted by: **The Lewin Group, with GDIT and Telligen**

This project was funded by the Centers for Medicare & Medicaid Services (CMS) under contract no. HHSM-500-2014-00033I. The statements contained in this report are solely those of the authors and do not necessarily reflect the views or policies of CMS. The Lewin Group assumes responsibility for the accuracy and completeness of the information contained in this report.

Prepared for:

Federal project officer

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

Submitted by:

The Lewin Group

3237 Airport Rd
La Crosse, WI 54603

Authors

Lewin: Amanda Markovitz, Soumita Lahiri, Elvedin Bijelic, Ian Morrall, Kassandra Fate, Colin Doyle, Niklaus Julius, Maxwell Mindock, Cody Hopewell, Stephen Poskitt, Lance Ford, Juan Castro, Hannah LeManske

Acknowledgments

The authors would also like to recognize contributions from additional team members:

Lewin: Rachel Henke, Brighita Negrusa, Mary Beth Schaefer, and Carol Simon

Telligen: Lindsay Kaatz and Christine LaRocca

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

This Word document was converted to a 508 PDF using CommonLook Office software (ver. 2.1.23.55). We also used CommonLook PDF software (ver. 5.20 32-bit) to make sure all tags were correct once the 508 PDF was created. The 508 PDF was checked for Accessibility with Adobe Acrobat Pro (ver. 26.001.216662).

Table of Contents

Appendix A: List of Acronyms and Glossary of Terms.....	1
Appendix B: Data and Methods.....	8
A. Secondary Data Sources	8
B. Study Population	11
C. Secondary Data Claims-Based Outcomes.....	20
D. Analytical Methodology	32
E. Savings to Medicare due to the CJR Model.....	47
Appendix C: Descriptive Statistics of the CJR Population	49
A. Descriptive Tables and Figures.....	49
Appendix D: Additional Findings	53
A. Model Impact.....	53
B. Medicare Program Savings.....	88

Appendix A: List of Acronyms and Glossary of Terms

Exhibit A-1: List of acronyms

Acronym	Meaning
ACH	Acute Care Hospital
ACO	Accountable Care Organization
ASC	Ambulatory Surgical Center
ATT	Average Treatment Effect on the Treated
BPCI	Bundled Payments for Care Improvement
BPCI Advanced	Bundled Payments for Care Improvement Advanced
C-APC	Comprehensive Ambulatory Payment Classification
CBSA	Core-Based Statistical Area
CCN	CMS Certification Number
CFR	Code of Federal Regulations
CI	Confidence Interval
CJR	Comprehensive Care for Joint Replacement
CMS	Centers for Medicare & Medicaid Services
COVID-19	Coronavirus Disease 2019
CPT	Current Procedural Terminology
CQS	Composite Quality Score
DDD	Difference-in-Difference-in-Differences
DiD	Difference-in-Differences
DME	Durable Medical Equipment
ED	Emergency Department
ESRD	End-Stage Renal Disease
FFS	Fee-for-Service
FR	Federal Register
FY	Fiscal Year
GPDC	Global and Professional Direct Contracting
HCC	Hierarchical Condition Category
HH	Home Health
HHA	Home Health Agency
HPES	High Post-Episode Spending

Acronym	Meaning
IP	Inpatient
IPO	Inpatient-Only
IPPS	Inpatient Prospective Payment System
IRF	Inpatient Rehabilitation Facility
LCI	Lower Confidence Interval Limit
LEJR	Lower Extremity Joint Replacement
LTCH	Long-Term Care Hospital
MA	Medicare Advantage
MCC	Major Complication or Comorbidity
MDM	Master Data Management
MDS	Minimum Data Set
MSA	Metropolitan Statistical Area
MS-DRG	Medicare Severity-Diagnosis Related Group
N	Number
NPRA	Net Payment Reconciliation Amount
NQF	National Quality Forum
OLS	Ordinary Least Squares
OP	Outpatient
OPPS	Outpatient Prospective Payment System
OT	Occupational Therapy
p	p-value
PAC	Post-Acute Care
PDP	Post-Discharge Period
PEP	Post-Episode Period
PGP	Physician Group Practice
PHE	Public Health Emergency
POS	Provider of Services
pp	Percentage Point
PT	Physical Therapy
PY	Performance Year
Q	Quarter (in a calendar year)
RIF	Research Identifiable File
SNF	Skilled Nursing Facility
SSP	Shared Savings Program

Acronym	Meaning
TAA	Total Ankle Arthroplasty
THA	Total Hip Arthroplasty
TKA	Total Knee Arthroplasty
UCI	Upper Confidence Interval Limit

Exhibit A-2: Glossary of terms

Term	Definition
Acute care hospital (ACH)	A health care facility that provides inpatient medical care and other related services for acute medical conditions or injuries.
Ambulatory surgical center (ASC)	A health care facility that provides surgical care to patients not requiring hospitalization or services exceeding 24 hours.
Anchor hospitalization or procedure	The inpatient hospitalization or outpatient LEJR procedure that triggers the start of the episode of care.
Baseline period	The baseline period used in the evaluation; the period of time that precedes the start of the CJR Model as a basis for comparison in the DiD statistical technique. The baseline period for our evaluation includes episodes that were initiated from 2012 to 2014 and that ended between April 1, 2012, and March 31, 2015.
Beneficiary notification requirement	Written notification that each participant hospital must provide to any Medicare beneficiary who meets the criteria for inclusion in the CJR Model.
Bundled payment model	A single payment or target price for the combined cost of eligible services and supplies – like treatments, tests, and procedures – provided during a defined episode of care. This payment or target price can cover multiple providers involved in the episode of care.
CJR collaborator	Medicare-enrolled providers and suppliers engaged in caring for CJR patients 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.
Composite quality score (CQS)	A summary score reflecting hospital performance and improvement on two LEJR-related quality measures (THA/TKA Complications measure and the HCAHPS Survey measure), as well as successful submission of THA/TKA patient-reported outcomes and limited risk variable data.
Effective discount percentage	The discount percentage, provided at the beginning of a performance period, serves as Medicare's portion of the savings. A 3% discount 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. The effective discount percentages are: 0% for "excellent" quality, 1.5% for "good" quality, and 3% for "acceptable" quality. The effective discount percentage for hospitals with "below acceptable" quality is also 3%, but hospitals with "below acceptable" quality are ineligible to receive reconciliation payments.
Elective episode	Episodes for inpatient LEJRs discharged under MS-DRG 469, MS-DRG 470, and OP LEJRs under CPT codes 27447 TKA and 27130 THA.

Term	Definition
Episode benchmark price	The episode benchmark price represents the expected episode payments if treatment patterns and patient mix do not change from previous historical spending for LEJR episodes. For Performance Year (PY) 6–PY 8, the episode benchmark price is based solely on regional amounts and includes both inpatient and outpatient TKA and THA, with separate benchmark prices for each MS/DRG and region. At reconciliation, benchmark prices are adjusted at the episode level for age bracket, hierarchical category condition counts, and dual eligibility for Medicaid. The product of the episode benchmark price and the effective discount percentage at reconciliation equals the quality-adjusted target price.
Episode of care	The set of services and supplies to treat a medical condition, for a defined length of time. A CJR episode of care was triggered by the admission of an eligible Medicare FFS beneficiary to a hospital paid under IPPS for an inpatient hospitalization that resulted in a discharge paid under one of the following MS-DRGs: 469 (major joint replacement or reattachment of lower extremity with MCC), 470 (major joint replacement or reattachment of lower extremity without MCC), 521 (hip replacement with principal diagnosis of hip fracture with MCC), or 522 (hip replacement with principal diagnosis of hip fracture without MCC). Beginning in PY 6, Total Knee Arthroplasty (CPT 27447) or Total Hip Arthroplasty (CPT 27130) procedures performed in the Hospital Outpatient Department (HOPD) setting and paid under the Outpatient Prospective Payment System (OPPS) also initiated a CJR episode. The episode ended 90 days after discharge from the anchor hospitalization or the outpatient procedure.
Extension period	PY 6–PY 8 of the CJR Model. The performance period of the CJR Model was extended by 3 years to evaluate design updates implemented by CMS in the 2021 Final Rule. The model extension applied to CJR participants in mandatory MSAs excluding hospitals with Section 401 rural status or low-volume status. CJR hospitals located in voluntary MSAs who previously opted in to continue their participation in the CJR Model were removed from the model at the start of the extension period.
Fracture episode	Episodes for LEJRs discharged under MS-DRG 521 and MS-DRG 522. These MS-DRGs only include hip fractures, and fractures of the knee or ankle are included in elective episode counts.
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.
High episode spending cap	The high episode spending cap prevented participant hospitals from being held responsible for catastrophic episode spending amounts that they could not reasonably have been expected to prevent by capping the costs for those episodes. In PY 1–PY 5, episodes were capped at 2 standard deviations above regional mean episode spending. In PY 6–PY 8, episodes were capped at the 99 th percentile of regional spending for each of the four MS-DRG categories.
Hospital referral region	Regional health care markets based on historical data of where most beneficiaries within a ZIP Code receive services for selected tertiary care, not including orthopedic services.
Inpatient-only (IPO) list	A list of procedures that are covered by Medicare only when provided in the inpatient setting.
Low volume	A hospital identified by CMS as having fewer than 20 LEJR episodes in total across the 3 historical years of data used to calculate the PY 1 CJR episode target prices.

Term	Definition
Mandatory MSA	Mandatory participation areas for the first 2 PYs included hospitals paid under the Medicare Inpatient Prospective Payment System (IPPS) in 67 MSAs, which consist of counties associated with a core urban area with a population of at least 50,000. As of February 1, 2018, 34 of the 67 areas remained mandatory MSAs and starting on October 1, 2021, hospitals in the 34 MSAs not designated as low volume or rural were designated as mandatory participants in CJR for PY 6–PY 8.
Metropolitan statistical area (MSA)	A geographical region that is associated with a core urban area and has a population of at least 50,000.
Normalization factor	The national mean of the target price for all episode types divided by the national mean of the risk-adjusted target price. Starting from PY 6, the normalization factor was applied during reconciliation, along with a retrospective market trend factor.
Net payment reconciliation amount (NPRA)	The total reconciliation target price amount minus the total performance period episode spending amount at the hospital-level, adjusted by stop gain or stop loss limits, if applicable.
Outpatient (OP) department	A hospital care setting for procedures not requiring admission or overnight stay 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 therapy services received by the patient from IRFs, SNFs, HHAs, LTCHs, or OP PTOT following a hospitalization.
Post-discharge period (PDP)	The 90-day period after discharge, which starts on the day of the anchor hospitalization discharge for inpatient episodes and the day of the procedure for outpatient episodes.
Quality-adjusted target price	A CJR hospital's target price at reconciliation after episode-level risk adjustment and application of a normalization factor, retrospective trend factor, and the hospital's effective quality discount factor.
Quality discount factors	A discount applied during calculation of reconciliation target prices for participant hospitals based on their CQS. In PY 6–PY 8, the quality discount is 0% for participant hospitals with “excellent” quality performance; 1.5% for participant hospitals with “good” quality performance; and 3% for participant hospitals with “below acceptable/acceptable” quality performance.
Reconciliation payment	A retrospective payment between Medicare and a CJR participant hospital at the end of a PY. If total FFS payments for its episodes during a PY are less than the aggregate quality-adjusted target price, Medicare makes a payment to a CJR participant hospital. If total FFS payments for a CJR participant hospital's episodes are more than its aggregate quality-adjusted target price, the hospital repays the difference to Medicare.
Related items and services	Episode-related items and services paid under Medicare Part A or Part B that are included in the bundle after exclusions are applied. These episode-related items and services include physicians' services, inpatient hospital services (including readmissions with certain exceptions discussed in the Final Rule), inpatient psychiatric facility services, LTCH services, IRF services, SNF services, HHA services, hospital outpatient services, outpatient therapy services, clinical laboratory services, DME, Part B drugs, and hospice.
Retrospective market trend factor	The regional/MS–DRG mean cost for episodes occurring during the PY divided by the regional/MS–DRG mean cost for episodes occurring during the target price base year.

Term	Definition
Risk adjustment	A statistical process to adjust claims-based outcomes to consider differences at the patient, episode, hospital, state, and MSA level that are related to the measures of interest. Without adequate risk adjustment, providers treating a sicker or more service-intensive patient mix would have worse outcomes than otherwise comparable providers serving healthier patients.
Rural (as defined by the CJR Model)	An IPPS hospital that is in a rural area as defined under § 412.64 of CMS Final Rule 42 CJR Part 510, is in a rural census tract defined under § 412.103(a)(1), or has reclassified as a rural hospital under § 412.103.
Site-neutral pricing	The concept of Medicare paying the same rate for the same service, irrespective of where the service is performed. In the CJR Model, site neutrality is applied in the extension period for inpatient LEJRs discharged under MS-DRG 470 and OP LEJRs discharged under CPT codes 27447 TKA and 27130 THA.
Stop-loss/Stop-gain limits	The stop-loss limit is the maximum amount a hospital will have to repay to CMS, and the stop-gain limit is the maximum amount that a hospital will receive from CMS as a reconciliation payment. They are based on a percentage of the quality-adjusted target price. The stop-loss and stop-gain limits are 20% in PY 6–PY 8.
Value-based care (VBC)	A healthcare delivery model that compensates providers based on patient health outcomes, quality of care, and efficiency.
Voluntary MSA	Beginning in 2018, the 67 original CJR MSAs were ranked by average historical wage-adjusted episode payment. CMS designated the bottom 33 MSAs as voluntary MSAs and gave participants in those MSAs a one-time opportunity to opt in to CJR Model through PY 5.

Appendix B: Data and Methods

In this appendix, we summarize the data and methods used to evaluate the CJR Model and generate the results presented in this Final Evaluation Report.

A. Secondary Data Sources

Exhibit B-1: Secondary data sources

Data source	Date range	Dataset contents	Use
BPCI-A programmatic participant data	<i>Intervention</i>	Identifies health care providers (hospitals, physicians, and physician practice groups) that are participating in the BPCI-A Model, the time period of participation, and the episodes for which they are participating.	Used to identify LEJR discharges in the control group that are assigned to BPCI-A participants for risk adjustment.
BPCI programmatic participant data	<i>Baseline and intervention</i>	Identifies health care providers (hospitals, PAC providers, physicians, and physician practice groups) that are participating in the BPCI initiative, the time period of participation, and the models and episodes for which they are participating.	Used to identify LEJR discharges that are assigned to BPCI participants for exclusion. Used to identify hospitals as past BPCI LEJR participants for risk adjustment.
CJR programmatic data	<i>Intervention</i>	List of CJR participant hospitals, as well as their PY 1, PY 2, PY 3, PY 4, PY 5, PY 6, PY 7, and PY 8 quality-adjusted target prices, reconciliation (NPRA), and hospital quality data.	Used to identify CJR participating hospitals, hospitals that continued mandatory participation in PY 3, their start and end dates in the CJR Model, their quality performance, and their reconciliation payments or repayment responsibility. Used total reconciliation payments and repayments to CMS to calculate savings to Medicare and investigate the distribution of NPRA.
FY Acute IPPS Final Rule data files	<i>FY 2016, FY 2018, FY 2021</i>	On an annual basis, CMS sets ACH IPPS payment rates. Data files include FY hospital-level information on provider identification number, bed count, medical residents per 1,000 beds, average daily census, disproportionate share hospital patient percentage, uncompensated care payment per claim, Medicare days as a percent of total inpatient days, and section 401 (rural) status.	Used to risk adjust for acute care IPPS hospital characteristics. Used to identify section 401 hospitals (rural designation) located in the control group to exclude from the PY 6–PY 8 sample.

Data source	Date range	Dataset contents	Use
CMS MDM	<i>Baseline and intervention</i>	Provider- and patient- 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 patients involved in Pioneer, Next Generation, and SSP 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 (SSP track 3, Comprehensive ESRD Care Model with downside risk, and Next Generation). For PY 6–PY 8, SSP Enhanced Track episodes were excluded.
Medicare Advantage penetration data	<i>Baseline and Intervention</i>	Enrollment counts (from CMS MA/Part D Contract Enrollment Data) for Medicare beneficiaries eligible for and enrolled in MA plans by county and month.	Used to create MA penetration rates for CJR MSAs over the time period of the evaluation, from evaluation baseline period through end of PY 8.
Medicare FFS beneficiary enrollment data	<i>Baseline and intervention</i>	Enrollment data (from Common Medicare Enrollment and Medicare Beneficiary Summary File) provide beneficiary Medicare Parts A and B eligibility information.	Enrollment data were used to confirm beneficiary eligibility and provide patient characteristics for analyses (for example, risk adjustment models). Enrollment data were used to measure the change in case-mix of CJR and control group patients between the baseline and the intervention periods.
Medicare FFS claims	<i>Baseline and intervention</i>	Medicare Parts A and B claims data (from monthly Medicare claims [TAP] files) provide claims for different services received during the anchor hospitalization and post-discharge period (for example, dates and types of service). A minimum of 3-month claims run out was used for episodes included in this report.	Claims were used to: 1) create the CJR episodes, describe service use, and create risk adjustment (for example, patient prior utilization, HCC score, COVID-19 diagnosis) and outcome variables (for example, unplanned readmissions, ED visits, and number of days or visits in each PAC setting); 2) identify TKA and THA procedures in the hospital outpatient departments and ASCs in CJR and control markets for descriptive analyses and create outpatient TKA and THA episodes; and 3) sample participants for primary data collection (patient survey, telephone interviews).

Data source	Date range	Dataset contents	Use
MDS 3.0 data	<i>Baseline and intervention</i>	The MDS is a comprehensive assessment instrument administered by nursing staff to all Medicare beneficiaries when they are admitted to a Medicare-certified SNF, at discharge, as well as on days 5, 14, 30, 60, 90, and quarterly, thereafter. The MDS collects information on patients' demographics, history and diagnoses, skin conditions, medications, care management, restraint use, preferences for routine and activities, and functional, sensory, cognitive, neuro or emotional, bladder, bowel, swallowing or nutritional, and pain status. A minimum 6 month run out of MDS data was used for episodes included in this report.	MDS data were used to identify patients who were in a SNF or long-term nursing facility during the 6 months preceding the episode which was used for risk-adjustment and to measure the change in case-mix of CJR patients and patients in the control group between the baseline and the intervention periods.
Medicare standardized payments	<i>Baseline and intervention</i>	Medicare standardized payments for 100% of Parts A and B claims received via the Chronic Condition Warehouse. Produced by a CMS contractor.	Used to create Medicare standardized paid amounts (Parts A and B) and allowed standardized payment amounts, including patient out-of-pocket amounts. Used to estimate the impacts of the CJR Model on total episode and service-level payments.
POS file	<i>December 2016</i>	Information on Medicare-approved facilities, including provider identification number, ownership status, size, medical school affiliation, and staffing.	Used to identify and characterize ACHs actively engaged in Medicare for risk adjustment and descriptive analyses.

Notes: ACO = Accountable Care Organization; BPCI = Bundled Payments for Care Improvement; BPCI-A = Bundled Payments for Care Improvement Advanced; CJR = Comprehensive Care for Joint Replacement; CMS = Centers for Medicare & Medicaid Services; COVID-19 = coronavirus disease 2019; FFS = fee-for-service; FY = fiscal year; HCC = Hierarchical Condition Category; ID = identification; IPPS = Inpatient Prospective Payment System; LEJR = lower extremity joint replacement; MDM = Master Data Management; MDS = Medicare Minimum Data Set 3.0; NPRA = net payment reconciliation amount; PAC = post-acute care; POS = provider of services; PY = performance year; SNF = skilled nursing facility; SSP = Shared Savings Program; THA = total hip arthroplasty; TKA = total knee arthroplasty.

B. Study Population

This section defines the CJR and control group populations, explains the sampling weights used in the mandatory analyses to account for differences in sampling probabilities, and outlines the additional eligibility criteria for hospitals and episodes.

1. *Defining the CJR and Control Group Populations*

At the start of the CJR Model in 2016, CMS selected MSAs eligible for CJR participation based on a stratified sampling methodology. CMS divided MSAs into eight strata based on historical wage-adjusted episode payments and population size. Within each stratum, MSAs were randomly selected to participate in the CJR Model (n = 67 MSAs). This design allowed for a control group of hospitals in MSAs that were eligible for CJR but not selected by CMS to participate in the CJR Model (n = 104 MSAs). These MSAs represented what would have happened in CJR-type markets if the model was never implemented (that is, the counterfactual). To be included in the CJR Model and in our analysis, hospitals had to be ACHs paid under the IPPS that performed LEJRs for Medicare patients.

In 2018, the 67 original CJR MSAs were ranked by average historical wage-adjusted episode payment, and CMS required the top 34 MSAs with the highest payments to continue participation in the model (mandatory MSAs) while giving hospitals in the bottom 33 MSAs a one-time opportunity to opt in (voluntary MSAs). This change reduced mandatory participation by about half. As discussed in **Chapter I: Introduction**, starting in PY 6, CMS removed all CJR participating hospitals in voluntary MSAs from the model. Additionally, CMS provided low-volume or rural CJR hospitals participating in mandatory MSAs a one-time choice to opt in to continued participation. After PY 6, these hospitals no longer participated in the model.

This report covers PY 6–PY 8 of the CJR Model, from October 1, 2021, to December 31, 2024. Our analyses focused on episodes from hospitals that were mandated to participate as of PY 6–PY 8. As such, we excluded rural and low-volume hospitals in the mandatory MSAs that previously could opt in to continue participation in CJR because they were no longer allowed to remain in the model in PY 6. Low-volume hospitals were defined as hospitals that had fewer than 20 episodes over a 3-year historical period (2012 to 2014), and we identified rural hospitals using the FY 2019 IPPS data (Section 401 hospitals). Starting in PY 6, a considerable number of CJR hospitals for which participation had been mandatory since the beginning of the model were no longer eligible to be in the model. Of the 395 mandatory CJR hospitals in PY 3–PY 5, 72 no longer participated in the model in PY 6 due to an updated rural designation status.

Exhibit B-2 provides the names of the CJR and control group MSAs included in the original CJR Model and in our analysis of PY 6–PY 8. MSAs that were included in our analysis of PY 6–PY 8 are indicated by an asterisk.

Exhibit B-2: CJR and control group MSAs included in the original model and PY 6–PY 8

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

CJR		Control	
CBSA ID	MSA name, state	CBSA ID	MSA name, state
31080	Los Angeles-Long Beach-Anaheim, CA*	20740	Eau Claire, WI
31180	Lubbock, TX*	22020	Fargo, ND-MN
31540	Madison, WI	22520	Florence-Muscle Shoals, AL*
32820	Memphis, TN-MS-AR*	22900	Fort Smith, AR-OK
33100	Miami-Fort Lauderdale-West Palm Beach, FL*	23060	Fort Wayne, IN
33340	Milwaukee-Waukesha-West Allis, WI	23420	Fresno, CA
33700	Modesto, CA	24340	Grand Rapids-Wyoming, MI
33740	Monroe, LA*	24580	Green Bay, WI
33860	Montgomery, AL*	24860	Greenville-Anderson-Mauldin, SC*
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*

CJR		Control	
CBSA ID	MSA name, state	CBSA ID	MSA name, state
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
		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

CJR		Control	
CBSA ID	MSA name, state	CBSA ID	MSA name, state
		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*

Notes: An asterisk indicates that the MSA had participating CJR hospitals in PY 6–PY 8 or was included in our control group for PY 6–PY 8. MSAs without an asterisk were not included in the analyses of PY 6–PY 8.

CBSA = core-based statistical area; CJR = Comprehensive Care for Joint Replacement; ID = identification; MSA = metropolitan statistical area; PY = performance year.

Source: CJR MSAs were pulled from the [CMS CJR Model web page](#). Information for control group MSAs provided by the Centers for Medicare & Medicaid Services.

2. Analytic MSA Sampling Weights and Construction of Control Group

In all analyses, unless otherwise noted, we used analytic MSA-level sampling weights that reflect both the stratified sampling of MSAs used in the design of the CJR Model and the voluntary–mandatory split of the model in 2018. The probability of an MSA being selected to participate in the original CJR Model varied across the strata, with CMS proportionally under-sampling MSAs in the lower average episode payment strata (strata 1, 2, 5, and 6) and over-sampling MSAs in higher average episode payment strata (strata 3, 4, 7, and 8). **Exhibit B-3** shows the count of CJR and control group MSAs by stratum and the proportion of MSAs in each stratum that made up the CJR and control groups.

Exhibit B-3: CMS' original stratified random sample of CJR MSAs

MSA population	MSA sampling stratum	MSA average episode payment	# MSAs eligible for sampling	CJR sample		Control group sample	
				# CJR MSAs	Percent of MSAs selected for CJR (%)	# Control group MSAs	Percent 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	39.2	104	60.8

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

Source: CJR evaluation team analysis of the CMS Final Rule Medicare Program Comprehensive Care for Joint Replacement Payment Model for ACHs Furnishing Lower Extremity Joint Replacement Services, 80 Fed. Reg. 73273 (November 24, 2015) (codified at 42 CFR 510).

CJR was implemented as a randomized control trial, so the control group began as the MSAs that were eligible but not selected to be mandatory CJR participants. In 2018, CMS chose 34 MSAs out of the 67 original MSAs to remain mandatory, based on ranked average historical MSA episode payments.

We then created MSA-specific analytic sampling weights to account for the exact details of CMS' stratified random sampling design and the selection of 34 mandatory MSAs as follows:

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

Step 2. We applied MSA-level weights to the 104 non-CJR MSAs based on the probability that the MSA was selected into the 34 mandatory CJR MSAs through the two-step selection process.

To construct the weights in Step 2, we first calculated the probabilities of the first-stage selection for each MSA, that is, the probability that the MSA was randomly selected to be in the original set of 67 CJR MSAs. These probabilities equaled the proportion of MSAs randomly selected for CJR from each MSA sampling stratum.

Next, we calculated the probabilities of the second-stage selection, that is, the probability that the MSA was selected among the 34 mandatory CJR MSAs given that it was selected in the first stage. Those second-stage selection probabilities were more complex to calculate because the MSAs for the 34 mandatory CJR MSAs were not selected randomly, so we could not rely on simple proportions.¹ We used an ATT analysis to evaluate the impact of CJR on mandatory hospitals; thus, we used the calculated probabilities to construct MSA-level weights such that the weighted control group was representative of the CJR group. For this reason, all CJR MSAs received a weight of 1. Specifically,

$$\text{Weight for mandatory CJR hospitals} = 1$$

$$\text{Weight for control group hospitals} =$$

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

These were MSA stratum-level weights so all control group hospitals in the same MSA had the same weight.

Exhibit B-4 shows the analytic weights calculated for control group MSAs and **Exhibit B-5** shows the full list of MSA-specific weights.

Exhibit B-4: Analytic sampling weights for control group MSAs

MSA sampling stratum	MSA	Weight
3	MSA-specific weights	Ranges from 0.00 to 0.73
4	All MSAs	1
7	MSA-specific weights	Ranges from 0.02 to 0.69
8	All MSAs	0.83

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

Source: CJR evaluation team analysis of the CMS Final Rule Medicare Program; Cancellation of Advancing Care Coordination Through Episode Payment and Cardiac Rehabilitation Incentive Payment Models; Changes to Comprehensive Care for Joint Replacement Payment Model: Extreme and Uncontrollable Circumstances Policy for the Comprehensive Care for Joint Replacement Payment Model, 82 Fed. Reg. 57066 (December 1, 2017) (codified at 42 CFR 510 and 42 CFR 512).

¹ The MSAs were selected by ranking the original 67 CJR MSAs by historical average episode payment and retaining the top half of the sample (that is, retaining the 34 MSAs with the highest historical average episode payment).

Exhibit B-5: Analytic sampling weights for control group MSAs

MSA sampling stratum	MSA	Weight
4	All MSAs	1.000
8	All MSAs	0.833
7	Birmingham-Hoover, AL	0.692
7	Cape Coral-Fort Myers, FL	0.692
7	Chattanooga, TN-GA	0.681
7	Columbia, SC	0.114
7	Greenville-Anderson-Mauldin, SC	0.689
3	Gulfport-Biloxi-Pascagoula, MS	0.727
3	Hattiesburg, MS	0.727
3	Huntsville, AL	0.707
3	Jonesboro, AR	0.727
7	Lexington-Fayette, KY	0.692
3	Lima, OH	0.727
7	Louisville/Jefferson County, KY-IN	0.508
3	Macon, GA	0.727
3	Manchester-Nashua, NH	0.000
7	North Port-Sarasota-Bradenton, FL	0.692
7	Philadelphia-Camden-Wilmington, PA-NJ-DE-MD	0.562
3	Punta Gorda, FL	0.727
7	Riverside-San Bernardino-Ontario, CA	0.015
7	Salt Lake City, UT	0.054
3	Waterloo-Cedar Falls, IA	0.726
7	Worcester, MA-CT	0.692
7	York-Hanover, PA	0.692

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

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).

3. LEJR Episode Definition

Initially, for both the CJR and control group populations, the beginning of an IP episode was triggered by an admission to a CJR participating or control group hospital (called an anchor hospitalization) with a resulting discharge in MS-DRG 469 or 470 (LEJR with MCC and LEJR without MCC, respectively). Starting in October 2020, CMS added two new MS-DRGs for LEJR due to hip fracture (521 with MCC and 522 without MCC) as episode triggers. Previously, hip fracture episodes were discharged under MS-DRGs 469 or 470 and were identified as having a hip fracture based on International Classification of Diseases diagnosis codes. The end of the episode is 90 days after the anchor hospitalization end date.

For OP LEJRs, the beginning of the episode is triggered by an LEJR performed in the OP department of a CJR participating or control group hospital, as identified in Part B institutional claims by CPT code 27447 for total knee arthroplasty or 27130 for total hip arthroplasty assigned to C-APC 5115 with status indicator “J1.” The end of the episode is 90 days after the OP procedure. OP episodes were added to the CJR Model starting in PY 6.

Identically for both IP and OP LEJR recipients, and consistent with CJR eligibility rules for PY 6–PY 8, Medicare beneficiaries who met and maintained the following eligibility criteria throughout the period were included in the analysis:

- Enrolled in Medicare Parts A and B;
- Medicare was the primary payer; and
- Not eligible for Medicare based on ESRD

Under CJR Model rules, episodes were canceled in the CJR Model and thus excluded from the evaluation’s analysis if:

- The patient no longer met the eligibility criteria described above;
- The patient was readmitted to a participating hospital during the episode and discharged under MS-DRG 469, 470, 521, or 522 (in which case the first episode is canceled and a new CJR episode begins);
- The patient died at any time during the episode period;
- The episodes started on or after July 1, 2017, and were prospectively assigned to a Next Generation ACO, an SSP ACO in track 3, or a Comprehensive ESRD Care Model ACO with downside risk;² or

² 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>

- The episodes were attributed to the BPCI initiative.³

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 patient information used to risk-adjust outcomes (for example, age, sex, and 6 months of Medicare FFS enrollment history prior to the LEJR hospital admission).

C. Secondary Data Claims-Based Outcomes

1. Measures of Impact on Payments, Utilization, and Quality

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

³ Episodes initiated at CJR participant hospitals could be attributed to a physician group practice participating in the BPCI initiative or to SNFs, IRFs, LTCHs, or HHAs participating in the BPCI Initiative Model 3.

Exhibit B-6: Claims-based outcome definitions⁴

Domain	Outcome name	Definition	Measurement period(s)	Eligible sample
Medicare payments	Total Medicare standardized allowed amounts per episode ⁵	The sum of Medicare payment and patient out-of-pocket amounts for related items and services covered by Medicare Part A and Part B performed during the LEJR hospitalization (anchor hospitalization) through the 90-day post-discharge period that are included in the episode. ⁶	Anchor hospitalization through 90-day post-discharge period	Patients who: 1) have a complete FFS enrollment history 6 months prior to the anchor hospitalization; 2) have consistent, reliable sex and age data (age <115); 3) maintain Parts A and B enrollment throughout the measurement period; 4) have a measurement period that ends on or before December 31, 2024; 5) have non-missing Medicare standardized allowed payment information for the episode; 6) have not died from any cause during the anchor hospitalization or 90-day post-discharge period.
	Total Medicare standardized paid amount per episode	The sum of Medicare Part A and B payments during the episode, not including any patient out-of-pocket amounts.	Anchor hospitalization through 90-day post-discharge period	Patients who: 1) have a complete FFS enrollment history 6 months prior to the anchor hospitalization; 2) have consistent, reliable sex and age data (age <115); 3) maintain Parts A and B enrollment throughout the measurement period; 4) have a measurement period that ends on or before December 31, 2024; 5) have non-missing Medicare standardized allowed payment information for the episode; 6) have not died from any cause during the anchor hospitalization or 90-day post-discharge period.

⁴ 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.

⁵ Standardized payments remove wage adjustments and other Medicare payment adjustments (for example, GME, IME, and DSH). Allowed amounts include beneficiary cost sharing.

⁶ Episode-related items and services paid under Medicare Part A or Part B, after exclusions are applied, including physician services, inpatient hospital services (including readmissions with certain exceptions discussed in the Final Rule), inpatient psychiatric facility services, 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 (continued)	Medicare standardized allowed amount for the anchor hospitalization per episode	The sum of Medicare payment and patient out-of-pocket amounts for the LEJR anchor hospitalization (MS-DRG 469, 470, 521, or 522 for IP episodes covered under Medicare Part A; CPT 27447 for outpatient TKA episodes and CPT 27130 for outpatient THA covered under Medicare Part B).	Anchor hospitalization	Patients who: 1) have a complete FFS enrollment history 6 months prior to the anchor hospitalization; 2) have consistent, reliable sex and age data (age <115); 3) maintain Parts A and B enrollment throughout the measurement period; 4) have a measurement period that ends on or before December 31, 2024; 5) have non-missing Medicare standardized allowed payment information for the episode; 6) have not died from any cause during the anchor hospitalization or 90-day post-discharge period.
	Other Medicare Part A standardized allowed amounts per episode, by service	The sum of Medicare payment and patient out-of-pocket amounts for all payments under Medicare Part A, excluding payments 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	Patients who: 1) have a complete FFS enrollment history 6 months prior to the anchor hospitalization; 2) have consistent, reliable sex and age data (age <115); 3) maintain Parts A and B enrollment throughout the measurement period; 4) have a measurement period that ends on or before December 31, 2024; 5) have non-missing Medicare standardized allowed payment information for the episode; 6) have not died from any cause during the anchor hospitalization or 90-day post-discharge period.
	Medicare standardized allowed amounts for HHA services per episode	The sum of Medicare payment and patient out-of-pocket amounts for HHA services covered under Medicare Part A or Part B HHA.	90-day post-discharge period	Patients who: 1) have a complete FFS enrollment history 6 months prior to the anchor hospitalization; 2) have consistent, reliable sex and age data (age <115); 3) maintain Parts A and B enrollment throughout the measurement period; 4) have a measurement period that ends on or before December 31, 2024; 5) have non-missing Medicare standardized allowed payment information for the episode; 6) have not died from any cause during the anchor hospitalization or 90-day post-discharge period.

Domain	Outcome name	Definition	Measurement period(s)	Eligible sample
Medicare payments (continued)	Other Medicare Part B standardized allowed amounts per episode	The sum of Medicare payment and patient out-of-pocket amounts for related items and services covered under Medicare Part B (except HHA services and payments for the LEJR if performed in the OP setting) 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	Patients who: 1) have a complete FFS enrollment history 6 months prior to the anchor hospitalization; 2) have consistent, reliable sex and age data (age <115); 3) maintain Parts A and B enrollment throughout the measurement period; 4) have a measurement period that ends on or before December 31, 2024; 5) have non-missing Medicare standardized allowed payment information for the episode; 6) have not died from any cause during the anchor hospitalization or 90-day post-discharge period.
	Medicare Part A SNF Standardized Allowed Amount	The sum of Medicare payment and patient out-of-pocket amounts for Part A health care services provided for SNF during the 90-day PDP.	90-day post-discharge period	Patients who: 1) have a complete FFS enrollment history 6 months prior to the anchor hospitalization; 2) have consistent, reliable sex and age data (age <115); 3) maintain Parts A and B enrollment throughout the measurement period; 4) have a measurement period that ends on or before December 31, 2024; 5) have non-missing Medicare standardized allowed payment information for the episode; 6) have not died from any cause during the anchor hospitalization or 90-day post-discharge period.
	Medicare Part A IRF Standardized Allowed Amount	The sum of Medicare payment and patient out-of-pocket amounts for Part A health care services provided for IRF during the 90-day PDP.	90-day post-discharge period	Patients who: 1) have a complete FFS enrollment history 6 months prior to the anchor hospitalization; 2) have consistent, reliable sex and age data (age <115); 3) maintain Parts A and B enrollment throughout the measurement period; 4) have a measurement period that ends on or before December 31, 2024; 5) have non-missing Medicare standardized allowed payment information for the episode; 6) have not died from any cause during the anchor hospitalization or 90-day post-discharge period.

Domain	Outcome name	Definition	Measurement period(s)	Eligible sample
Medicare payments (continued)	Medicare Part A Readmissions Standardized Allowed Amount	The sum of Medicare payment and patient out-of-pocket amounts for Part A health care services provided for readmissions during the 90-day PDP.	90-day post-discharge period	Patients who: 1) have a complete FFS enrollment history 6 months prior to the anchor hospitalization; 2) have consistent, reliable sex and age data (age <115); 3) maintain Parts A and B enrollment throughout the measurement period; 4) have a measurement period that ends on or before December 31, 2024; 5) have non-missing Medicare standardized allowed payment information for the episode; 6) have not died from any cause during the anchor hospitalization or 90-day post-discharge period.
	Medicare standardized allowed amounts for services provided in the 30 days post-episode per episode	The sum of Medicare payment and patient 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	Patients who: 1) have a complete FFS enrollment history 6 months prior to the anchor hospitalization; 2) have consistent, reliable sex and age data (age <115); 3) maintain Parts A and B enrollment throughout the measurement period; 4) have a measurement period that ends on or before January 30, 2025; 5) have non-missing Medicare standardized allowed payment information for the episode; 6) have not died from any cause during the anchor hospitalization or 90-day post-discharge period.
Utilization: First PAC destination	First discharge to IRF	The percentage of all episodes with patients initially discharged to an IRF. The first PAC setting is an IRF (a freestanding facility or a distinct unit within an acute hospital) if admission to the IRF occurred within the first 5 days of hospital discharge and no other PAC use occurred prior to IRF admission. If the patient is directly transferred to another ACH after the anchor hospitalization, then the first PAC setting was defined within 5 days of the transfer discharge.	1 st to 5 th day after discharge from the anchor or transfer hospitalization	Patients who: 1) have a complete FFS enrollment history 6 months prior to the anchor hospitalization; 2) have consistent, reliable sex and age data (age <115); 3) maintain Parts A and B enrollment throughout the measurement period; 4) have a measurement period that ends on or before December 31, 2024; 5) have non-missing Medicare standardized allowed payment information for the episode; 6) have not died from any cause during the anchor hospitalization or 90-day post-discharge period.

Domain	Outcome name	Definition	Measurement period(s)	Eligible sample
Utilization: First PAC destination (continued)	First discharge to SNF	The percentage of all episodes with patients initially discharged to a SNF. The first PAC setting is a SNF if admission to the SNF occurred within the first 5 days of hospital discharge and no other PAC use occurred prior to SNF admission. If the patient is directly transferred to another ACH after the anchor hospitalization, then the first PAC setting was defined within 5 days of the transfer discharge.	1 st to 5 th day after discharge from the anchor or transfer hospitalization	Patients who: 1) have a complete FFS enrollment history 6 months prior to the anchor hospitalization; 2) have consistent, reliable sex and age data (age <115); 3) maintain Parts A and B enrollment throughout the measurement period; 4) have a measurement period that ends on or before December 31, 2024; 5) have non-missing Medicare standardized allowed payment information for the episode; 6) have not died from any cause during the anchor hospitalization or 90-day post-discharge period.
	First discharge to HHA	The percentage of all episodes with patients 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 patient is directly transferred to another ACH after the anchor hospitalization, then the first PAC setting was defined within 14 days of the transfer discharge.	1 st to 14 th day after discharge from the anchor or transfer hospitalization	Patients who: 1) have a complete FFS enrollment history 6 months prior to the anchor hospitalization; 2) have consistent, reliable sex and age data (age <115); 3) maintain Parts A and B enrollment throughout the measurement period; 4) have a measurement period that ends on or before December 31, 2024; 5) have non-missing Medicare standardized allowed payment information for the episode; 6) have not died from any cause during the anchor hospitalization or 90-day post-discharge period.
	First discharge to OP PT/OT	The percentage of all episodes with patients initially discharged to an outpatient physical therapy or occupational therapy setting. The first PAC setting is an OP PT/OT if there was not an admission to either SNF or IRF within 5 days after discharge, not an admission to an HHA within 14 days of hospital discharge, but there was a use of OP PT/OT within days 2 to 14 after discharge. If the patient is directly transferred to another ACH after the anchor hospitalization, then the first PAC setting was defined within 14 days of the transfer discharge.	1 st to 14 th day after discharge from the anchor or transfer hospitalization	Patients who: 1) have a complete FFS enrollment history 6 months prior to the anchor hospitalization; 2) have consistent, reliable sex and age data (age <115); 3) maintain Parts A and B enrollment throughout the measurement period; 4) have a measurement period that ends on or before December 31, 2024; 5) have non-missing Medicare standardized allowed payment information for the episode; 6) have not died from any cause during the anchor hospitalization or 90-day post-discharge period.

Domain	Outcome name	Definition	Measurement period(s)	Eligible sample
Utilization: First PAC destination (continued)	First discharge to other	The percentage of all episodes with patients initially discharged to a setting other than SNF, IRF, HHA, or OP PT/OT. The first PAC setting is other if there was not an admission to either SNF or IRF within 5 days after discharge, not an admission to a HHA within 14 days of hospital discharge, not an admission to a OP PT/OT within days 2 to 14 after discharge of the discharge, but there was an admission to hospice, an IPPS, an LTCH, or another IP setting within 5 days of hospital discharge. If the patient is directly transferred to another ACH after the anchor hospitalization, then the first PAC setting was defined within 14 days of the transfer discharge.	1 st to 14 th day after discharge from the anchor or transfer hospitalization	Patients who: 1) have a complete FFS enrollment history 6 months prior to the anchor hospitalization; 2) have consistent, reliable sex and age data (age <115); 3) maintain Parts A and B enrollment throughout the measurement period; 4) have a measurement period that ends on or before December 31, 2024; 5) have non-missing Medicare standardized allowed payment information for the episode; 6) have not died from any cause during the anchor hospitalization or 90-day post-discharge period.
	First discharge home without HHA or OP PT/OT	The percent of all episodes with patients initially discharged to home without HHA or OP PT/OT services. The first PAC setting is home without HHA or OP PT/OT if the patient is not admitted to a SNF or IRF within 5 days of hospital discharge, is not admitted to a HHA within 14 days of hospital discharge, does not utilize OP PT/OT services on days 2 to 14 after discharge, and does not use an admitted to a different location within 5 days after discharge. If the patient is directly transferred to another ACH after the anchor hospitalization, then the first PAC setting was defined within 14 days of the transfer discharge.	1 st to 14 th day after discharge from the anchor or transfer hospitalization	Patients who: 1) have a complete FFS enrollment history 6 months prior to the anchor hospitalization; 2) have consistent, reliable sex and age data (age <115); 3) maintain Parts A and B enrollment throughout the measurement period; 4) have a measurement period that ends on or before December 31, 2024; 5) have non-missing Medicare standardized allowed payment information for the episode; 6) have not died from any cause during the anchor hospitalization or 90-day post-discharge period.

Domain	Outcome name	Definition	Measurement period(s)	Eligible sample
Utilization	Any HHA use	The percent of all episodes with patients using any HHA services during the 90-day post-discharge period, as indicated by non-zero Medicare payment and patient out-of-pocket amounts for HHA services covered under Medicare Part A or Part B.	90-day post-discharge period	Patients who: 1) have a complete FFS enrollment history 6 months prior to the anchor hospitalization; 2) have consistent, reliable sex and age data (age <115); 3) maintain Parts A and B enrollment throughout the measurement period; 4) have a measurement period that ends on or before December 31, 2024; 5) have non-missing Medicare standardized allowed payment information for the episode; 6) have not died from any cause during the anchor hospitalization or 90-day post-discharge period.
	Number of IRF days	The average number of IRF days of care during the 90-day post-discharge period.	90-day post-discharge period	Patients who: 1) have a complete FFS enrollment history 6 months prior to the anchor hospitalization; 2) have consistent, reliable sex and age data (age <115); 3) maintain Parts A and B enrollment throughout the measurement period; 4) have a measurement period that ends on or before December 31, 2024; 5) have at least 1 IRF day during this period; 6) have non-missing Medicare standardized allowed payment information for the episode; 7) have not died from any cause during the anchor hospitalization or 90-day post-discharge period.
	Number of SNF days	The average number of SNF days of care during the 90-day post-discharge period.	90-day post-discharge period	Patients who: 1) have a complete FFS enrollment history 6 months prior to the anchor hospitalization; 2) have consistent, reliable sex and age data (age <115); 3) maintain Parts A and B enrollment throughout the measurement period; 4) have a measurement period that ends on or before December 31, 2024; 5) have at least 1 SNF day during this period; 6) have non-missing Medicare standardized allowed payment information for the episode; 7) have not died from any cause during the anchor hospitalization or 90-day post-discharge period.

Domain	Outcome name	Definition	Measurement period(s)	Eligible sample
Utilization (continued)	Number of outpatient PT or OT visits	The average number of outpatient PT or OT visits during the 90-day post-discharge period.	90-day post-discharge period	Patients who: 1) have a complete FFS enrollment history 6 months prior to the anchor hospitalization; 2) have consistent, reliable sex and age data (age <115); 3) maintain Parts A and B enrollment throughout the measurement period; 4) have a measurement period that ends on or before December 31, 2024; 5) have at least one outpatient PT or OT visit during this period; 6) have non-missing Medicare standardized allowed payment information for the episode; 7) have not died from any cause during the anchor hospitalization or 90-day post-discharge period.
Quality	Unplanned readmission rate	The proportion of episodes with one or more unplanned readmissions for any eligible condition. ⁷ Following these specifications, we excluded planned admissions, based on Agency for Healthcare Research and Quality Clinical Classification System Procedure and Diagnoses codes.	90-day post-discharge period	Patients who: 1) have a complete FFS enrollment history 6 months prior to the anchor hospitalization; 2) have consistent, reliable sex and age data (age <115); 3) maintain Parts A and B enrollment throughout the measurement period; 4) have a measurement period that ends on or before December 31, 2024; 5) are discharged from the anchor hospitalization in accordance with medical advice; 6) have non-missing Medicare standardized allowed payment information for the episode ; 7) have not died from any cause during the anchor hospitalization or 90-day post-discharge period.

⁷ Updated specification documents were released by CMS in March 2019 for the unplanned readmission measure, and the measure was revised accordingly. Available at: <https://www.qualitynet.org/inpatient/measures/readmission/methodology>

Domain	Outcome name	Definition	Measurement period(s)	Eligible sample
Quality (continued)	ED visit rate	The proportion of episodes with one or more ED visits during the 90-day post-discharge period for which the patient required medical treatment but was not admitted to the hospital. Eligible ED visits are outpatient claims with a code indicating the patient used the ED but was not admitted to the hospital.	90-day post-discharge period	Patients who: 1) have a complete FFS enrollment history 6 months prior to the anchor hospitalization; 2) have consistent, reliable sex and age data (age <115); 3) maintain Parts A and B enrollment throughout the measurement period; 4) have a measurement period that ends on or before December 31, 2024; 5) are discharged from the anchor hospitalization in accordance with medical advice; 6) have non-missing Medicare standardized allowed payment information for the episode; 7) have not died from any cause during the anchor hospitalization or 90-day post-discharge period.
	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 patients at CJR participant and control group hospitals that would have been identified as episodes if they had not died during the episode of care. Patients who: 1) have a complete FFS enrollment history 6 months prior to the anchor hospitalization; 2) have consistent, reliable sex and age data (age <115); 3) maintain Parts A and B enrollment throughout the measurement period; 4) have not received hospice care in the 6 months prior to admission; 5) have a measurement period that ends on or before December 31, 2024; 6) are discharged from the anchor hospitalization in accordance with medical advice; 7) have non-missing Medicare standardized allowed payment information for the episode.

Domain	Outcome name	Definition	Measurement period(s)	Eligible sample
Quality (continued)	Incidence of any complications	The proportion of episodes with incidence (during the anchor hospitalization or a readmission) of: acute myocardial infarction, pneumonia, or sepsis or septicemia within the 7-day PDP; or surgical site bleeding or pulmonary embolism within the 30-day PDP; or mechanical complications, periprosthetic joint infection, or wound infection within the 90-day PDP. This measure was based on specifications for the NQF-endorsed THA or TKA complications measure (NQF measure 1550) for elective LEJRs.⁸ Death in the 30 days after discharge is part of the technical definition but is not included in our analysis because patients who died during the anchor hospitalization or in the 90-day PDP are excluded from the CJR Model.	90-day post-discharge period	Patients who: 1) have a complete FFS enrollment history 6 months prior to the anchor hospitalization; 2) have consistent, reliable sex and age data (age <115); 3) maintain Parts A and B enrollment throughout the measurement period; 4) have a measurement period that ends on or before December 31, 2024; 5) are discharged from the anchor hospitalization in accordance with medical advice; 6) have non-missing Medicare standardized allowed payment information for the episode; 7) have not died from any cause during the anchor hospitalization or 90-day post-discharge period.

Notes: ACH = acute care hospital; CJR = Comprehensive Care for Joint Replacement; CMS = Centers for Medicare & Medicaid Services; CPT = current procedural terminology; DME = durable medical equipment; ED = emergency department; FFS = fee-for-service; HHA = home health agency; IPPS = Inpatient Prospective Payment System; IRF = inpatient rehabilitation facility; LEJR = lower extremity joint replacement; LTCH = long-term care hospital; NQF = National Quality Forum; OP = outpatient; 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.

⁸ Updated specification documents were released by CMS in March 2019 for the THA or TKA complications measure, and the measure was revised accordingly. Available at: <https://www.qualitynet.org/inpatient/measures/complication/methodology>

2. Generalizability of CJR Results

In our evaluation of the CJR Model, we aimed to assess how generalizable the model's impacted populations were to the broader U.S. healthcare landscape. We examined the extent to which LEJR episodes from PY 6–PY 8 in CJR-participating hospitals resembled those in non-participating areas. **Exhibit B-7** lists the episode and patient characteristics from claims and enrollment data that we monitored.

Exhibit B-7: Measure of episode and patient characteristics

Type of characteristic	Measure name or description
Episode characteristic	Total LEJRs
	Hospital OP LEJR
	Elective LEJR
	Fracture LEJR
Patient characteristic	Dual Status
	Diabetes
	Obesity
	Dementia
	Age
	HCC Score

Notes: HCC = Hierarchical Condition Category; LEJR = lower extremity joint replacement; OP = outpatient.

Source: Episode and patient characteristic measures were constructed from Medicare FFS claims and beneficiary enrollment data.

In **Chapter II: Overview of the CJR Population**, we presented focused results comparing CJR and control episodes with national hospital LEJRs, where national hospital LEJRs reflected episodes as if CJR Model rules had been applied nationwide. In [Appendix C: Descriptive Statistics of the CJR Population](#), we provided a more detailed generalizability analysis that incorporated multiple groups formed through the sequential exclusion of LEJRs. **Exhibit B-8** presents a detailed explanation of the sequential exclusion criteria, which are applied from top to bottom in the exhibit. Analytic sampling weights were not used for generalizability analyses because the sample extended beyond the CJR and control groups, and weights were not available for the additional MSAs. Applying CJR-specific weights to an expanded sample could introduce bias.

Exhibit B-8: Subgroup descriptions for generalizability analysis

Subgroup title	Subgroup description
National LEJRs	All Medicare FFS LEJRs performed across the United States.
Eligible LEJRs	LEJRs for patients who meet CMS model criteria: Medicare A and B beneficiaries only, alive throughout the episode period, not eligible for ESRD, non-IPPS, with no readmissions under MS-DRG 469, 470, 521, or 522, and no transfers
Non-ACO LEJRs	LEJRs that were not excluded due to model overlap with SSP or NextGen ACOs
Data eligible LEJRs	LEJRs for patients with at least 6 months of FFS enrollment history, reliable sex and age data, maintained Parts A and B enrollment throughout the measurement period, and complete Medicare payment information
Hospital LEJRs	LEJRs that were performed at hospitals (not performed in an ASC)
MSA LEJRs	LEJRs that occurred in MSAs (not in micro statistical or rural areas)
Model MSA LEJRs	LEJRs in a mandatory CJR or control MSA
Model hospital LEJRs	LEJRs in a mandatory CJR or control hospital
CJR episodes	LEJRs in a mandatory CJR hospital

Notes: ACO = Accountable Care Organization; ASC = ambulatory surgery center; CJR = Comprehensive Care for Joint Replacement; CMS = Centers for Medicare & Medicaid Services; ESRD = end-stage renal disease; FFS = fee-for-service; IPPS = Inpatient Prospective Payment System; LEJR = lower extremity joint replacement; MSA = metropolitan statistical area; MS-DRG = Medicare Severity Diagnosis Related Group; SSP = Shared Savings Program.

Source: Episode and patient characteristic measures were constructed from Medicare FFS claims and beneficiary enrollment data.

D. Analytical Methodology

To control for both observed and unobserved differences and to isolate the impact of the CJR Model on outcomes, we employed a DiD regression approach using the control group created by the model's randomization of CJR MSAs and supplemented by risk adjustment.

1. DiD Estimator

The DiD approach quantified the impact of the CJR Model by comparing changes in outcomes between the baseline and the intervention period of interest (PY 6–PY 8) for the CJR population and the control group population. One of the main advantages of this approach is that it can isolate the effect of unobserved characteristics of treatment and control groups that are time invariant.⁹

⁹ Although 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 encompassed episodes that started between January 1, 2012, and December 31, 2014, and ended between April 1, 2012, and March 31, 2015.

b. Intervention Period

The intervention period for this annual report followed the model timeline for PY 6–PY 8 and was specified as episodes that ended between October 1, 2021, and December 31, 2024.¹⁰

c. Exclusion of Performance Years 1–5

In this annual report, we were focused on studying the impact of the CJR Model during PY 6–PY 8. Thus, the methodology did not need to account for artifacts of the CJR Model prior to PY 6 if they were no longer relevant in PY 6.

d. Primary Regression Specification

The DiD model used an outcome measure, Y , and estimated the differential change in Y for patients receiving care from CJR participant hospitals between the baseline and the intervention period relative to that same change for patients receiving care from hospitals in the control group.

As an illustrative example of the DiD approach, we defined:

- $Y_{i,k,t}$ as the outcome for the i^{th} episode with an LEJR at hospital k in period t ($t = 1$ during the CJR PY 6–PY 8 intervention quarters and 0 otherwise)
- $CJR_{i,k}$ as an indicator that takes the value of 1 if the i^{th} episode was initiated by a CJR participant hospital k and takes the value of 0 otherwise.
- $X_{i,k,t}$ as hospital, geographic, and patient characteristics in period t .
- $E[Y|t, CJR, X]$ as the expected value of outcome measure Y conditional on values of t , CJR , and X .

The DiD estimator can then be represented by:

$$DiD = [E(Y | t = 1, CJR = 1, X) - (E(Y | t = 0, CJR = 1, X))] - [E(Y | t = 1, CJR = 0, X) - (E(Y | t = 0, CJR = 0, X))] \quad (1)$$

¹⁰ Final Rule issued by CMS in the FR on April 29, 2021, extending the CJR Model: <https://www.federalregister.gov/documents/2020/02/24/2020-03434/medicare-program-comprehensive-care-for-joint-replacement-model-three-year-extension-and-changes-to>.

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

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

- 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 captured the relative differences in outcomes between CJR and control group episodes.
- Coefficient b_3 determined the differential in outcome Y experienced by patients receiving services from CJR hospitals during the CJR intervention period relative to control group episodes in the intervention period and represented the DiD estimator.
- The vector of coefficients B measured the marginal effects of risk factors (X) on the outcome variable.

Finally, to calculate the DiD estimate for outcome measures that were risk-adjusted with nonlinear models, we used the regression model's coefficient estimates to calculate each of the four conditional expectations that make up the DiD estimator in Equation 1. For all DiD models, we assessed statistical significance at the 10% level but did not impose strict cutoffs of p-values when interpreting results. For example, given the context and magnitude of impact estimates, we consider some results with p-values just above 0.1 as important findings and others with p-values below 0.1 as borderline. Given the design of the CJR Model, which involved randomly sampling MSAs to determine participation, we accounted for clustering at the MSA level in the estimation of our standard errors in all regression models.

Additionally, we conducted various tests of whether the CJR and control populations were on parallel trends in the baseline period. See [Section D.1.g](#) for a description of these methodologies and [Appendix D: Additional Findings](#) for results.

e. Secondary Regression Specification

In addition to the primary regression specification, we estimated a secondary regression for each outcome. These specifications were identical to those described above, but instead of combining PY 6–PY 8 into a single “intervention” term in the DiD design, we estimate a separate effect for each performance year. No other elements of the regression changed. Results are in [Appendix D: Additional Findings](#).

Specifically, to illustrate the calculation of the PY specific DiD, consider the linear model listed below:

$$Y_{i,k,t} = b_0 + b_1 \cdot PY6 + b_2 \cdot PY7 + b_3 \cdot PY8 + b_4 \cdot CJR_{i,k} + b_5 \cdot CJR_{i,k} \cdot PY6 + b_6 \cdot CJR_{i,k} \cdot PY7 + b_7 \cdot CJR_{i,k} \cdot PY8 + X_{i,k,t}' \cdot B + u_{i,k,t} \quad (3)$$

- The value of coefficient b_1 captured aggregate factors that could cause changes in outcome Y in PY 6 relative to the baseline period that are common across CJR and control group episodes.
- The value of coefficient b_2 captured aggregate factors that could cause changes in outcome Y in PY 7 relative to the baseline period that are common across CJR and control group episodes.
- The value of coefficient b_3 captured aggregate factors that could cause changes in outcome Y in PY 8 relative to the baseline period that are common across CJR and control group episodes.
- Coefficient b_4 captured the relative differences in outcomes between CJR and control group episodes.
- Coefficient b_5 determined the differential in outcome Y experienced by patients receiving services from CJR hospitals during PY 6 relative to control group episodes in PY 6 and represented the DiD estimator for PY 6.
- Coefficient b_6 determined the differential in outcome Y experienced by patients receiving services from CJR hospitals during PY 7 relative to control group episodes in PY 7 and represented the DiD estimator for PY 7.
- Coefficient b_7 determined the differential in outcome Y experienced by patients receiving services from CJR hospitals during PY 8 relative to control group episodes in PY 8 and represented the DiD estimator for PY 8. The vector of coefficients B measured the differential effects of risk factors (X) on the outcome variable.

f. Risk Adjustment to Control for Differences in Patient Demographics and Clinical Risk Factors

In the DiD models, we controlled for potential differences in patient demographics, clinical characteristics observed before hospitalization, and provider characteristics (represented by $X_{i,k,t}$ in Equation 2 above; **Exhibit B-9**). Some demographic factors included age categories, sex, age and sex interactions, Medicaid eligibility status, and disability status. We risk-adjusted all outcomes for procedure type (hip or knee) and presence or absence of an MCC, defined by MS-DRGs.¹¹ To control for participation in other Medicare initiatives, we used an indicator variable that indicated whether the patient was in the SSP ACO, Pioneer ACO Model, or Next Generation ACO

¹¹ We made two modifications to DRGs for this process: (1) We back-code DRGs that occurred prior to the separation of elective and fracture DRGs to match what they would have been under this change, and (2) we code all OP procedures as MS-DRG 470.

Model during the episode.¹² To control for prior health conditions, we used HCC indicators for the 12 months preceding the anchor hospitalization, as well as indicators for obesity, diabetes, hypertension, and tobacco use, generated from the claims data in the CCW.¹³ To further control for case-mix differences, we included measures of the following types of prior care use: acute care IPPS hospital, SNF, IRF, hospice, other Part A IP, HHA, custodial nursing facility, and others. We used the Long-Term Care MDS 3.0 data to create a measure of prior custodial nursing facility use in the 6 months prior to the episode. The collection of this assessment data was temporarily paused during most of the COVID-19 Public Health Emergency; thus, we imputed seemingly missing values of this covariate during that time period. We also included COVID-19-related risk-adjusters. To address patient-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).

We also controlled for provider characteristics that might have been related to the outcomes of interest, such as hospital bed count, ownership type, previous BPCI initiative LEJR experience, and previous BPCI initiative experience in a clinical episode other than LEJR. In October 2018, the BPCI Advanced initiative began. This Innovation Center model also included LEJR as a clinical episode and aimed to reduce payments while maintaining or improving quality. CJR participant hospitals could not participate in the BPCI Advanced initiative for LEJR clinical episodes; however, hospitals and surgeons in the control group could participate. To account for contamination in our control group by this other Innovation Center model, we included indicator variables that identify control group LEJR episodes performed by surgeons or at hospitals participating in the BPCI Advanced Model.

¹² Patients with episodes during or after July 2017 that were aligned with the SSP track 3, Next Generation ACO, or the Comprehensive ESRD Care Model were excluded from the CJR Model.

¹³ The CMS-HCC Model is a prospective risk-adjustment model used by CMS to adjust Medicare Part C capitation payments for beneficiary health spending risk. The model adjusts for demographic and clinical characteristics. The clinical component of the model uses diagnoses from qualifying services grouped into numerous HCC indicators.

Exhibit B-9: Predictive risk factors used to risk-adjust claims-based outcomes for impact analyses

Domain	Variables
Characteristics of the procedure	Anchor MS-DRG
	Procedure type (hip or knee)
Patient demographics and enrollment	Age (under 65, 65–79, 80+ years)
	Sex
	Medicaid status
	Disability status at enrollment in Medicare (not ESRD)
	Attribution to Shared Savings Program, Pioneer ACO Model, or Next Generation ACO Models during the Comprehensive Care for Joint Replacement episode
	Other demographic characteristics
Prior health conditions	CMS-HCC version 21 indicators from qualifying services and diagnoses from claims and data for 12 months preceding the anchor hospitalization ^a
	Obesity indicator from the CCW
	Diabetes indicator from the CCW
	Hypertension indicator from the CCW
	Tobacco use indicator from the CCW
Prior use	Prior care use (any acute care IP, ED visits, IRF, SNF, HHA, hospice, other Part A IP, LTCH, and custodial nursing facility service) variables used in risk adjustment varied by model^b • Binary indicators for any care use in the 6 months preceding the start of the episode • Binary indicators for any care use in the 1 month preceding the start of the episode • Number of days of care use in the 6 months preceding the start of the episode
Geography	State indicators
COVID-19	Patient-level COVID-19 diagnosis in the 30 days prior to the anchor hospitalization
	Patient-level COVID-19 diagnosis during the anchor hospitalization
Hospital provider characteristics	Bed count
	For-profit status
	BPCI LEJR experience
	BPCI experience in a clinical episode other than LEJR
	LEJR performed by surgeons or at hospitals actively participating in the BPCI-A Model for LEJR clinical episodes (control group only; episode-level variable)

Notes: ACO = Accountable Care Organization; BPCI = Bundled Payments for Care Improvement; BPCI-A = Bundled Payments for Care Improvement Advanced; CCW = Chronic Condition Warehouse; CMS = Centers for Medicare & Medicaid Services; COVID-19 = coronavirus disease 2019; ED = emergency department; ESRD = end-stage renal disease; HCC = Hierarchical Condition Category; HHA = home health agency; IP = inpatient;

IRF = inpatient rehabilitation facility; LEJR = lower extremity joint replacement; LTCH = long-term care hospital; MS-DRG = Medicare severity diagnosis related group; SNF = skilled nursing facility.

- ^a The HCC indicators in the risk adjustment model included: sepsis, different types of cancer, diabetes, obesity, malnutrition, rheumatoid arthritis, coagulation defects, dementia, drug or alcohol dependence, mood disorder, Parkinson's disease, seizure disorders, cardio-respiratory failure, congestive heart failure, angina, heart arrhythmias, stroke, vascular disease, chronic obstructive pulmonary disease, macular degeneration, kidney disease, and renal failure. For hip fracture only analyses, we also include indicators for metastatic cancer and acute leukemia, lung and other severe cancers, intestinal obstruction/perforation, inflammatory bowel disease, severe hematological disorders, disorders of immunity, dementia with complications, drug/alcohol psychosis, schizophrenia, spinal cord disorders/injuries, acute myocardial infarction, cerebral hemorrhage, hemiplegia/hemiparesis, fibrosis of lung and other chronic lung disorders, aspiration and specified bacterial pneumonias, chronic kidney disease severe (stage 4), pressure ulcer of skin with partial thickness skin loss, pressure pre-ulcer skin changes of unspecified stage, major head injury, vertebral fractures without spinal cord injury, and artificial openings for feeding or elimination. (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.)
- ^b The specification for each prior use variable varied for each outcome. The binary 6-month indicators were used for: SNF payment, IRF payment, HHA payment, Part B payment, unplanned readmissions, ED use, number of SNF days, and first discharge setting. The binary 1-month indicators were used for: complications and mortality. The indicators for number of days in the past 6 months were used for: total episode payment, readmissions payment, 30-day post-episode payment, number of IRF days, number of HHA visits, any HHA visits, and number of outpatient PT or OT visits.

Source: Risk adjustment variables were constructed from Medicare FFS claims and beneficiary enrollment data, December 2016 provider of services, FY 2016 CMS Annual Inpatient Prospective Payment System, CMS Master Data Management, BPCI initiative participant list, and BPCI-A initiative participant list.

g. Parallel Trends

One critical assumption for an unbiased DiD estimate is that the average outcomes for both groups would have followed parallel trends in the intervention period had the policy not been implemented. As with any DiD design, the parallel trend assumption concerns a counterfactual outcome path (what would have occurred in the absence of the policy) which cannot be observed once the policy is implemented. As such, we tested to see if CJR and the control group outcomes followed parallel trends for the outcome of interest during the baseline period, as a proxy for whether they likely would have followed parallel trends in the intervention period.

We evaluated the parallel trends assumption three ways, with each way testing whether the CJR and control group outcomes followed parallel trends during the baseline period. In the first two methods, we estimated episode-level models for each outcome using baseline data and used both a joint test of equality on discrete time periods and a linear trends test to study whether there was statistical evidence that the groups exhibited parallel trends in the baseline period. We considered outcomes to fail these statistical parallel trends tests if we rejected the null hypothesis of seemingly parallel trends at the 10% significance level. In the third method, we descriptively estimated a "Hypothetical DiD," which took into account any potential differential pre-trends between CJR and control. We then compared it with the actual estimated PY 6–PY 8 DiD to gauge how large an effect potential differential pre-trends would have had on the DiD impact estimate. These findings help provide interpretation and context for our standard PY 6–PY 8 DiD estimates.

i. Statistical Tests of Parallel Trends

For the joint test of equality on discrete time periods, we report the p-value of an F-test that tested whether the differential between the CJR and control groups was jointly equal across discrete 4-quarter time periods. We included dummy variables for each of the 3 baseline years, interaction terms between the CJR group indicator and each of the year dummies, along with all the risk-adjustment variables that we included in the DiD models.

The joint F-test model was:

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

where:

- $Y_{i,k,t}$ was the outcome for the i^{th} episode with an LEJR at hospital k in the baseline period in year t ,
- $Year_{i,t}$ was an indicator that took the value of 1 if the i^{th} episode was initiated during year t of the baseline period and took the value of 0 otherwise,
- $CJR_{i,k}$ was an indicator that took the value of 1 if the i^{th} episode was initiated by a CJR participant hospital k and took the value of 0 otherwise,
- $X_{i,k}$ were hospital, geographic, and patient characteristics in the baseline period,

and the test was:

$$H_0: b_4 = b_5 = b_6$$

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

For the linear trends test, we report the p-value of a linear slope coefficient of the quarterly difference between the CJR and control group. We included a quarterly indicator; interaction term between the CJR group indicator and the quarterly indicator, along with all the risk-adjustment variables that we included in the DiD models.

The linear test model was:

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

where:

- $Y_{i,k,t}$ was the outcome for the i^{th} episode with an LEJR at hospital k in the baseline period in quarter t ,
- $Quarter_{i,t}$ was an indicator that took the value of 1 if the i^{th} episode was initiated during quarter t of the baseline period and took the value of 0 otherwise,
- $CJR_{i,k}$ was an indicator that took the value of 1 if the i^{th} episode was initiated by a CJR participant hospital k and took the value of 0 otherwise,

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

and the test was:

$$H_0: b_3 = 0$$

$$H_1: b_3 \neq 0$$

ii. Hypothetical Difference-in-Difference Method

In addition to using the joint test and linear trend test to determine whether the CJR and control populations were on differential trends in the baseline period, we conducted a third parallel trend analysis to examine the extent to which possible deviations in parallel trends could possibly influence our PY 6–PY 8 combined impact estimates, as reported in **Chapter III: Impact of the Model**. While the calculations of the hypothetical DiD impact estimates were objective, their interpretation was fundamentally subjective. In this section, we describe our approach, our findings, and ultimately our interpretation, acknowledging that the interpretation may differ for different readers.

The first step in calculating the hypothetical impact estimates involved running a two-stage least squares difference-in-difference regression which controlled for differential baseline trends between CJR and control. This regression was almost identical to that of our standard difference-in-difference equation described in **Section D.1.d**, with the primary difference being that we allowed for both CJR and control to have a trend in the baseline period.

The model entails two first-stage equations:

$$\begin{aligned} \text{Control Quarter}_{i,k,t} &= b_0 + b_1 \cdot \text{Modified Control Quarter}_{i,k,t} + b_2 \cdot \text{Modified CJR Quarter}_{i,k,t} + b_3 \\ &\quad * \text{CJR}_k + b_4 * \text{After}_t + b_5 * \text{After}_t * \text{CJR}_k + X_{i,k}' \cdot B + u_{i,k,t} \end{aligned}$$

and

$$\begin{aligned} \text{CJR Quarter}_{i,k,t} &= b_0 + b_1 \cdot \text{Modified Control Quarter}_{i,k,t} + b_2 \cdot \text{Modified CJR Quarter}_{i,k,t} + b_3 \\ &\quad * \text{CJR}_k + b_4 * \text{After}_t + b_5 * \text{After}_t * \text{CJR}_k + X_{i,k}' \cdot B + u_{i,k,t} \end{aligned}$$

where:

- $\text{Control Quarter}_{i,k,t}$ was the quarter for the i^{th} episode with an LEJR at hospital k in the control group in quarter t . Observations at hospital k in the CJR group have a value of 0 for all quarters.
- $\text{CJR Quarter}_{i,k,t}$ was the quarter for the i^{th} episode with an LEJR at hospital k in the CJR group in quarter t . Observations at hospital k in the control group have a value of 0 for all quarters.
- $\text{Modified Control Quarter}_{i,k,t}$ was the quarter for the i^{th} episode with an LEJR at hospital k in the control group in the baseline period in quarter t . All observations at hospital k in the

CJR group and, additionally, all observations in a quarter in the intervention period have a value of 0.

- *Modified CJR Quarter*_{*i,k,t*} was the quarter for the *i*th episode with an LEJR at hospital *k* in the CJR group in the baseline period in quarter *t*. All observations at hospital *k* in the control group and, additionally, all observations in a quarter in the intervention period have a value of 0.
- *CJR_k*, *After_t*, and *After_t * CJR_k + X_{i,k}* are the standard DiD covariates, and *X_{i,k}* are hospital, geographic, and patient characteristics. See **Section D.1.d** for more details.

Each of these first stage equations generate predicted values of the quarter variable for control or CJR observations, respectively. Let these predicted variables be labeled as

*Control Quarter*_{*i,k,t*} and *CJR Quarter*_{*i,k,t*}

Without loss of generality, we will focus on *Control Quarter*_{*i,k,t*}. Given the estimating equation above, *Control Quarter*_{*i,k,t*} is predicted for all observations, regardless of whether the observation is for CJR or control, and regardless of whether it is baseline or intervention. Specifically:

- Observations in control in the baseline: *Control Quarter*_{*i,k,t*} is almost perfectly predicted for all control observations in the baseline period due to the inclusion of *Modified Control Quarter*_{*i,k,t*} variable.¹⁴
- Observations in CJR (both baseline and intervention): As *Control Quarter*_{*i,k,t*} is equal to 0 for all CJR observations, this variable is almost perfectly predicted (to equal 0) for all CJR observations due to the inclusion of the *CJR_k* variable.
- Observations in control in the intervention: All control observations in the intervention period are almost predicted to equal the average value of *Control Quarter*_{*i,k,t*} in the intervention period due to the inclusion of the *After_t* variable.

The second stage of the model is then estimated by:

$$Outcome_{i,k,t} = b_0 + b_1 \cdot \widehat{Control\ Quarter}_{i,k,t} + b_2 \cdot \widehat{CJR\ Quarter}_{i,k,t} + b_3 \cdot CJR_k + b_4 \cdot After_t + b_5 \cdot After_t \cdot CJR_k + X_{i,k}' \cdot B + u_{i,k,t}$$

where:

- *Control Quarter*_{*i,k,t*} and *CJR Quarter*_{*i,k,t*} are the predicted variables from the first-stage model and all other variables are as previously specified,
- The differential linear trend estimate is taken as the difference between *b₂* and *b₁*, and
- *b₅* is the hypothetical DiD estimate.

¹⁴ Note, these observations are not truly perfectly predicted due to extremely slight deviations caused by risk-adjustment. If we were to run these models without risk-adjustment, these observations would be perfectly predicted in the first stage. Throughout this chapter, we use the word “almost” to refer to this type of prediction behavior.

The advantage of the two-stage linear regression method is that the baseline trends for CJR and control are estimated in the same regression as the hypothetical DiD. This leads the estimated coefficients for all the risk-adjustment included in $X_{i,k}$ to much more closely resemble those estimated in our standard DiD as they are estimated using the exact same sample as compared to a standard parallel trends model that only used baseline data.

Importantly, the coefficients on the baseline trends for CJR and control (b_1 and b_2 , respectively) are not impacted by any trends in the intervention period. Again without loss of generality, this is because while the intervention control observations of $\widehat{Control\ Quarter}_{i,k,t}$ are not equal to zero, the fact they are all approximately equal to the average quarter value for control observations in the intervention means that they are almost perfectly collinear with the observations that dictate the value the coefficient on $After_t$ (b_5). That is, the temporal effects of control observations in the intervention period are captured by $After_t$, leaving only the trending temporal effects of control observations in the baseline period to influence b_1 (the coefficient $\widehat{Control\ Quarter}_{i,k,t}$ on). These temporal effects are not perfectly captured by the constant (b_0), as that only captures the average effect in the baseline, not an effect that is allowed to trend over the baseline quarters.

We interpreted and presented these results holistically. We paid close attention to both the magnitude and statistical significance of the PY 6–PY 8 DiD estimate, the magnitude and statistical significance of the differential trend coefficient, and the results for the Joint F-Test and Linear Test (shown in **Exhibit D-1**). While using various pieces of information from multiple results to assist in a single interpretation of the effect of the CJR Model was an inherently subjective process, we have applied the process as consistently across samples and outcomes as possible.

We classified outcomes into four groupings with the following guidelines. Note, while we used these as guidelines, we stress the subjective nature of this analysis, and we present all intermediate pieces of information that were used in our interpretation in this section. For interpretation purposes, we did not consider standard p-value cutoffs, for example, <0.10 as “meaningful” as strictly binding. For example, we interpreted a p-value of 0.12 nearly identically to a p-value of 0.09.

- Very High Trust: We believe we can interpret the DiD estimates with relative certainty for these outcomes.
 - a. Outcomes that did not statistically fail the parallel trends tests and where the differences between the baseline trends of the CJR and control samples were of small magnitude.
 - b. Outcomes that did not statistically fail the parallel trends tests and had sufficiently large standard errors, such that we did not believe we could extrapolate the trends in any meaningful way.
- High Trust: Outcomes that did potentially statistically fail the parallel trends tests at the 10% level, but for which the differential trend was small in magnitude. We believe we could interpret the DiD estimates with relative certainty for these outcomes.

- Low Trust: Outcomes that did not statistically fail the parallel trends test but did have sufficiently large differential baseline trends. For these outcomes, we believe additional caveats may be warranted in the interpretation of the DiD estimate.
- Very Low Trust: Outcomes that did statistically fail parallel trends tests and had sufficiently large differential pre-trends. We believe strong additional caveats were warranted in the interpretation of the DiD estimates for these outcomes.

h. Impact on Patients Dually Eligible for Medicare and Medicaid

Our analysis of the differential impact of the CJR Model on patients who were dually eligible for Medicare and Medicaid was based on the DiD methodology described above. In essence, we estimated the impact of the CJR Model on patients who were dually eligible and the impact on patients who were not dually eligible and then estimated the difference between those two estimates. In practice, all these estimations took place in a single DDD regression, subject to the constraint that coefficients on risk-adjustment variables were the same for both dually eligible and not dually eligible episodes. These risk-adjustment variables are different from the set we use for the main impact section and are listed in **Exhibit B-10**.

Exhibit B-10: Predictive risk factors used to risk-adjust claims-based outcomes for impact on patients dually eligible for Medicare and Medicaid analyses

Domain	Variables
Characteristics of the procedure	Anchor MS-DRG
	Procedure type (hip or knee)
Patient demographics and enrollment	Age (under 65, 65–79, 80+ years)
	Sex
	Medicaid status
	Disability status at enrollment in Medicare (not ESRD)
	Attribution to Shared Savings Program, Pioneer ACO Model, or Next Generation ACO Models during the CJR episode
	Other demographic characteristics
Prior health conditions	Total HCC score in 12 months prior to the episode
	Dementia indicator ^a
	Hypertension indicator ^a
Prior use	Prior care use (any acute care IP, ED visits, IRF, SNF, HHA, hospice, other Part A IP, LTCH, and custodial nursing facility service) variables used in risk adjustment varied by model^b Binary indicators for any care use in the 6 months preceding the start of the episode Binary indicators for any care use in the 1 month preceding the start of the episode Number of days of care use in the 6 months preceding the start of the episode
Geography	State indicators

Domain	Variables
COVID-19	Patient-level COVID-19 diagnosis in the 30 days prior to the anchor hospitalization
	Patient-level COVID-19 diagnosis during the anchor hospitalization
Hospital provider characteristics	Bed count
	For-profit status
	BPCI LEJR experience
	BPCI experience in a clinical episode other than LEJR
	LEJR performed by surgeons or at hospitals actively participating in the BPCI-A Model for LEJR clinical episodes (control group only; episode-level variable)

Notes: ACO = Accountable Care Organization; BPCI = Bundled Payments for Care Improvement; BPCI-A = Bundled Payments for Care Improvement Advanced; CCW = Chronic Condition Warehouse; CMS = Centers for Medicare & Medicaid Services; COVID-19 = coronavirus disease 2019; ED = emergency department; ESRD = end-stage renal disease; HCC = Hierarchical Condition Category; HHA = home health agency; IP = inpatient; IRF = inpatient rehabilitation facility; LEJR = lower extremity joint replacement; MS-DRG = Medicare Severity-Diagnosis Related Group; SNF = skilled nursing facility.

^a Combined indicator created using a combination of CMS HCC version 21 indicators and claims-based indicators from the CCW. If either indicator was present, the condition was assumed to be present.

^b The specification for each prior use variable varied for each outcome. The binary 6-month indicators were used for: SNF payment, IRF payment, HHA payment, Part B payment, unplanned readmissions, ED use, number of SNF days, and first discharge setting. The binary 1-month indicators were used for: complications and mortality. The indicators for number of days in the past 6 months were used for: total episode payment, readmissions payment, 30-day post-episode payment, number of IRF days, number of HHA visits, any HHA visits, and number of outpatient PT or OT visits.

Source: Risk adjustment variables were constructed from Medicare FFS claims and beneficiary enrollment data, December 2016 provider of services, FY 2016 CMS Annual Inpatient Prospective Payment System, CMS Master Data Management, BPCI initiative participant list, and BPCI-A initiative participant list.

In previous reports, we only estimated differential impacts for either all LEJRs or elective LEJRs because of limited sample sizes. With the addition of PY 8 data, we had a sufficiently large sample size to also estimate differential impacts specifically for fracture LEJRs this analysis.

In general, the statistical model to estimate this DDD regression was:

$$Y_{i,k,t} = b_0 + b_1 \cdot t + b_2 \cdot CJR_{i,k} + b_3 \cdot G_{i,t} + b_4 \cdot G_{i,t} \cdot t + b_5 \cdot CJR_{i,k} \cdot t + b_6 \cdot CJR_{i,k} \cdot G_{i,t} \cdot t + X'_{i,k,t} \cdot B + u_{i,k,t}$$

where,

- $Y_{i,k,t}$ was the outcome for the i^{th} episode with an LEJR at hospital k in period t ($t = 1$ during the CJR PY 6–PY 8 intervention quarters and zero otherwise)
- $CJR_{i,k}$ was an indicator that takes the value of 1 if the i^{th} episode was initiated by a CJR participant hospital k and takes the value of 0 otherwise
- $G_{i,t}$ was an indicator that takes the value of 1 if the patient for the i^{th} episode was dually eligible and takes the value of 0 if the patient was not dually eligible

- $X_{i,k,t}$ was a vector containing hospital, geographic, and patient characteristics in period t ¹⁵
- The value of coefficient b_1 captured aggregate factors that could cause changes in outcome Y in the intervention period relative to the baseline period that are common across CJR and control group episodes.
- Coefficient b_2 captured the relative differences in outcomes between CJR and control group episodes.
- Coefficient b_3 captured the relative differences in outcomes between patients with dual eligibility and patients without dual eligibility
- Coefficient b_4 captured aggregated factors that could cause changes in outcome Y for patients with dual eligibility in the intervention period relative to the baseline period that are common across CJR and control group episodes.
- Coefficient b_5 determined the differential in outcome Y experienced by patients receiving services from CJR hospitals during the CJR intervention period relative to control group episodes in the intervention period and represented the DiD estimator.
- Coefficient b_6 determined the differential in outcome Y experienced by patients with dual eligibility receiving services from CJR hospitals during the CJR intervention period relative to control group episodes in the intervention period and represented the DDD estimator.
- The vector of coefficients B measured the effects of risk factors (X) on the outcome variable.

2. Model Types

We used a variety of models including logistic, Poisson, multinomial logit, OLS regressions, and two-part models (**Exhibit B-11**). Models were chosen depending on the type and characteristics of the outcome measure. For example, logistic models were estimated for the discrete quality outcomes (that is, all claims-based quality-of-care measures). A multinomial logit model was applied to the first discharge setting. Due to insufficient sample size, observations with a first discharge setting of “other” were omitted from the multinomial logit analyses of first discharge setting. OLS models were estimated for the continuous total number of days or visits measures as well as total episode payments and Part B payments. We used two-part models for payment outcomes where a considerable number of individuals had zero payments for the particular outcome.

¹⁵ Note, we used a reduced set of health condition indicators when estimating the impact on patients dually eligible for Medicaid due to issues with over and perfect prediction.

Exhibit B-11: Outcomes by model type

Model type	Outcomes
OLS	Total episode payments
	Part B payments
	Number of IRF days
	Number of SNF days
	Number of OP PT or OT Visits
Two-part models (Probit or OLS)	Readmission payments
	IRF payments
	SNF payments
	HHA payments
	30-day post-episode payments
Multinomial logistic	First post-acute discharge was to IRF
	First post-acute discharge was to SNF
	First post-acute discharge was to HHA
	First post-acute discharge was to OP PT/OT
	Discharge to home without HH or OP PT/OT ¹
Logistic	Any HHA visits
	Unplanned readmission
	ED visit
	Complications
	All-cause mortality

Notes: ED = emergency department; HHA = home health agency; IRF = inpatient rehabilitation facility; OLS = ordinary least squares; OP = outpatient; OT = occupational therapy; PT = physical therapy; SNF = skilled nursing facility.

¹ Some episodes meet the requirements of being classified as “First post-acute discharge was to Other”. However, as these observations make up less than 1% of observations, these observations are dropped from the multinomial logistic regression.

We used estimates from the multivariate regression models to construct model-predicted outcomes (sometimes labeled “risk-adjusted mean outcomes”) under two scenarios (baseline and PY 6–PY 8 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 patients, we used the same reference population of patients to calculate predicted outcomes for CJR and control group episodes. The reference population used for this report was all CJR patients during the baseline and intervention period.

E. Savings to Medicare due to the CJR Model

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

$$\text{Medicare savings} = \text{Change in nonstandardized paid amounts} - \text{Net Reconciliation payments}$$

1. Change in Nonstandardized Paid Amounts

To best capture the actual amount of payments sent and received by CMS, we used nonstandardized paid amounts for our analyses of the savings to Medicare. We calculated the change in nonstandardized paid amounts using estimates from a DiD model of per-episode standardized paid amounts. The DiD estimates were multiplied by –1 and converted to nonstandardized paid amounts using the ratio of nonstandardized-to-standardized Medicare paid amounts from CJR intervention episodes (**Exhibit B-12**). This method produced a per-episode estimate of the change in nonstandardized paid amounts, with positive values equating to reductions in per-episode spending. We calculated the total change in nonstandardized paid amounts by multiplying the per-episode estimate by the total number of PY 6–PY 8 CJR episodes.

Exhibit B-12: Ratios of nonstandardized-to-standardized Medicare paid amounts over time

Time period	Mandatory hospitals
Baseline	1.038
PY 6	1.059
PY 7	1.047
PY 8	1.085

Notes: The ratio was calculated as the average nonstandardized (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.

CJR = Comprehensive Care for Joint Replacement; PY = performance year.

Source: CJR evaluation team analysis of Medicare claims and enrollment data for episodes initiated in 2012 through 2014 that ended between April 2012 and March 2015 (baseline) and episodes that ended between October 1, 2021, and December 31, 2024 (PY 6–PY 8 intervention).

Note that in this report, we estimated Medicare savings only for PY 6–PY 8. The Medicare savings results for PY 1–PY 5 that are reported in the [CJR Fifth Annual Report](#) comprehensively describe Medicare savings over the first 5 PYs of the model, and so we cite those results when appropriate.

2. Reconciliation Payments

We defined net reconciliation payments as the total reconciliation payments made to CJR participants by Medicare minus repayments from CJR participants to Medicare. Net reconciliation payments could be positive or negative. In the program literature, they are often referred to as NPRA. The CMS CJR payment contractor provided this data. We calculated

reconciliation payments per episode by dividing total reconciliation payments by the total number of CJR episodes.

In analyzing the distribution of NPRA to hospitals, we ordered hospitals by their NPRA in PY 6–PY 8 and grouped the ordered hospitals into deciles, each comprising approximately 10% of the total hospital participants. For each decile, we calculated the total NPRA received by hospitals in the decile and average NPRA across hospitals in the decile. We previously performed a similar analysis for total NPRA in PY 1–PY 5 in the [CJR Fifth Annual Report](#). We then conducted exploratory analyses aimed at identifying the characteristics of hospitals that consistently received reconciliation payments or owed substantial repayments.

3. Target Prices

We also analyzed target prices. Target prices in the CJR Model are set at the episode level in PY 6–PY 8, considering the type of procedure, the hospital region, market trends, and patient characteristics. The CMS CJR payment contractor provided this data. We calculated the average target price within each MS-DRG by averaging the target price across all episodes of that MS-DRG.¹⁶ We calculated the average target price overall by averaging across all episodes.

4. Considerations

In the estimation of Medicare program savings, we dropped some episodes that resulted in reconciliation payments or repayments to CMS from the estimation sample, primarily due to our requirement that patients have a complete FFS enrollment history for 6 months prior to the anchor hospitalization. We explored this restriction in the [CJR Fourth Annual Report](#). We did not extrapolate estimated payment reductions to these “missing” episodes. Thus, our estimates of total Medicare program savings are slightly conservative; if the missing episodes also had payment reductions due to the CJR Model, our estimates of Medicare program savings are underestimates.

In figure notes where we report Medicare program savings estimates, we noted that the ranges for net savings are ranges based on the 90% CI. We specified this because technically the net savings estimate are not CIs given that we treat net reconciliation payments as non-random. This is because if gross reductions in episode spending had been different, net reconciliation payments would also have been different. The reported ranges for net savings estimates incorporate the uncertainty in our estimate of gross episode spending reductions, but do not incorporate a modelled relationship between gross episode spending reductions and NPRA.

¹⁶ In PY 6–PY 8, outpatient episodes were included in the CJR Model. Outpatient episodes are not classified into MS-DRGs, but are instead classified using HCPCS codes. In this analysis we classified outpatient episodes as the appropriate MS-DRG (for instance, an elective outpatient hip replacement would be classified as MS-DRG 470).

Appendix C: Descriptive Statistics of the CJR Population

In this appendix, we present the descriptive tables focusing on the CJR and control population in this Final Evaluation Report.

A. Descriptive Tables and Figures

Exhibit C-1: Elective and fracture LEJRs for CJR and control populations in PY 6–PY 8

Performance year	Quarter	CJR		Control	
		Elective LEJRs (N)	Fracture LEJRs (N)	Elective LEJRs (N)	Fracture LEJRs (N)
PY 6	Q4 2021	8,732	1,176	6,865	883
	Q1 2022	9,041	1,262	7,165	870
	Q2 2022	10,014	1,290	7,560	985
	Q3 2022	9,592	1,292	7,445	830
	Q4 2022	9,694	1,233	7,290	886
PY 7	Q1 2023	9,479	1,287	7,100	880
	Q2 2023	10,746	1,232	8,204	840
	Q3 2023	10,306	1,271	7,687	880
	Q4 2023	9,868	1,226	7,538	884
PY 8	Q1 2024	9,532	1,278	7,146	889
	Q2 2024	10,720	1,362	7,986	888
	Q3 2024	9,869	1,266	7,207	855
	Q4 2024	9,426	1,175	7,011	848

Notes: Elective LEJRs include both inpatient and outpatient episodes. Fracture LEJRs include only inpatient episodes.

CJR = Comprehensive Care for Joint Replacement Model; LEJR = lower extremity joint replacement;
N = number; PY = performance year; Q = quarter.

Source: CJR evaluation team analysis of Medicare claims and enrollment data for episodes that ended between October 1, 2021, and December 31, 2024 (PY 6–PY 8 intervention).

Exhibit C-2: Proportions of outpatient elective LEJRs for CJR and control populations in PY 6–PY 8

Elective LEJRs	Performance year	CJR			Control		
		Total LEJRs (N)	OP LEJRs (N)	Proportion (%)	Total LEJRs (N)	OP LEJRs (N)	Proportion (%)
THA	PY 6	17,301	11,605	67.1	12,577	9,205	73.2
	PY 7	14,498	10,913	75.3	10,377	8,133	78.4
	PY 8	14,381	11,315	78.7	10,294	8,330	80.9
TKA	PY 6	29,642	20,337	68.6	23,649	17,783	75.2
	PY 7	25,809	19,722	76.4	20,075	16,256	81.0
	PY 8	25,081	20,462	81.6	18,977	15,928	83.9

Notes: Total ankle arthroplasty (TAA) episodes are excluded from these results.

CJR = Comprehensive Care for Joint Replacement Model; LEJR = lower extremity joint replacement; N = number; OP = outpatient; PY = performance year; THA = total hip arthroplasty; TKA = total knee arthroplasty.

Source: CJR evaluation team analysis of Medicare claims and enrollment data for episodes that ended between October 1, 2021, and December 31, 2024 (PY 6–PY 8 intervention).

Exhibit C-3: Demographic, clinical, and episode characteristics for national and CJR patient populations in PY 6–PY 8

Outcome	National LEJRs	Eligible LEJRs	Non-ACO LEJRs	Data eligible LEJRs	Hospital LEJRs	MSA LEJRs	Model MSA LEJRs	Model hospital LEJRs	CJR episodes
Total LEJRs	1,652,690	1,548,086	1,497,141	1,436,214	1,291,447	1,168,449	511,479	322,500	143,319
Hospital OP LEJR	60.6%	62.7%	63.8%	63.8%	71.0%	71.5%	69.0%	68.7%	65.9%
ASC LEJR	9.6%	10.1%	10.2%	10.1%	0.0%	0.0%	0.0%	0.0%	0.0%
Elective LEJR with MCC or TAA	1.1%	1.0%	0.9%	0.9%	1.0%	1.1%	1.1%	1.0%	1.1%
Elective LEJR without MCC	87.1%	89.2%	89.5%	89.4%	88.2%	88.6%	88.1%	87.9%	87.5%
Fracture LEJR with MCC	3.0%	2.1%	2.0%	2.1%	2.3%	2.2%	2.3%	2.3%	2.4%
Fracture LEJR without MCC	8.7%	7.7%	7.5%	7.6%	8.5%	8.1%	8.5%	8.7%	9.1%
Dual status	7.5%	6.8%	6.9%	6.5%	7.0%	6.8%	6.6%	6.1%	7.2%
Diabetes	23.6%	23.4%	23.5%	23.8%	24.6%	24.5%	26.0%	26.0%	27.1%
Obesity	31.7%	32.4%	32.6%	33.2%	33.8%	34.1%	35.3%	34.7%	34.1%
Dementia	4.9%	4.1%	4.1%	4.2%	4.6%	4.5%	4.9%	5.0%	5.2%
Age	74.2	74.1	74.1	74.4	74.7	74.7	74.9	74.9	75.1
HCC score	1.40	1.36	1.36	1.38	1.44	1.44	1.49	1.49	1.51

Notes: National LEJRs include all Medicare FFS LEJRs performed across the United States. Eligible LEJRs include LEJRs for patients who meet CMS model criteria: Medicare A and B beneficiaries only, alive throughout the episode period, not eligible for ESRD, non-IPPS, with no readmissions under MS-DRG 469, 470, 521, or 522, and no transfers. Non-ACO LEJRs are LEJRs that were not excluded due to model overlap with SSP or NextGen ACOs. Data eligible LEJRs are those for patients with at least 6 months of FFS enrollment history, reliable sex and age data, maintained Parts A and B enrollment throughout the measurement period, and complete Medicare payment information. Hospital LEJRs are those that were not performed in an ASC. MSA LEJRs are those not in a micro statistical or rural areas. Model MSA LEJRs are those in a mandatory CJR or control MSA. Model hospital LEJRs are those in a mandatory CJR or control hospital. CJR episodes are those in a mandatory CJR hospital.

ACO = Accountable Care Organization; ASC = ambulatory surgical center; CJR = Comprehensive Care for Joint Replacement Model; ESRD = end-stage renal disease; FFS = fee-for-service; HCC = Hierarchical Condition Category; IPPS = inpatient prospective payment system; LEJR = lower extremity joint replacement; MCC = major complication or comorbidity; MSA = metropolitan statistical area; MS-DRG = Medicare Severity Diagnosis-Related Group; OP = outpatient; PY = performance year; SSP = Shared Savings Program; TAA = total ankle arthroplasty.

Source: CJR evaluation team analysis of Medicare claims and enrollment data for episodes that ended between October 1, 2021, and December 31, 2024 (PY 6–PY 8 intervention).

Exhibit C-4: MA penetration rates in CJR and control MSAs over the evaluation

Notes: MA and total Medicare beneficiary counts are from CMS public use files and are aggregated to the MSA level. Averages are calculated at the monthly level using the total LEJR volume and MSA analytic sampling weights. Month is measured by the month of the episode start date.

CJR = Comprehensive Care for Joint Replacement Model; LEJR = lower extremity joint replacement; MA = Medicare Advantage; MSA = metropolitan statistical area; PY = performance year.

Source: CJR evaluation team analysis of Medicare claims and enrollment data for episodes initiated after January 1, 2012, and that ended before December 31, 2024, and publicly available [Medicare Advantage and Enrollment Data](#).

Appendix D: Additional Findings

In this appendix, we present detailed results for the DiD impacts on our claims-based outcomes and Medicare program savings. Brief descriptions of the methodology are provided when applicable. For a full description of our data and methods please see [Appendix B: Data and Methods](#).

A. Model Impact

The results presented in this section correspond to the findings presented in **Chapter III: Impact of the Model**.

1. Parallel Trends

Exhibit D-1: Linear and joint tests of parallel trends for payment, utilization, and quality metrics, baseline, all LEJR episodes, fracture episodes, and elective episodes

Domain	Measure	All LEJR		Elective		Fracture	
		Joint test	Linear test	Joint test	Linear test	Joint test	Linear test
Payments	<i>Total episode payments</i>	p = 0.801	p = 0.903	p = 0.678	p = 0.726	p = 0.857	p = 0.325
	<i>SNF payments</i>	p = 0.749	p = 0.728	p = 0.684	p = 0.880	p = 0.419	p = 0.278
	<i>IRF payments</i>	p = 0.056	p = 0.454	p = 0.037	p = 0.571	p = 0.103	p = 0.084
	<i>HH payments</i>	p = 0.003	p = 0.001	p = 0.008	p = 0.003	p = 0.394	p = 0.459
	<i>Readmission payments</i>	p = 0.314	p = 0.276	p = 0.224	p = 0.152	p = 0.886	p = 0.692
	<i>Anchor payments</i>	p = 0.009	p = 0.140	p = 0.008	p = 0.108	p = 0.920	p = 0.636
	<i>Other Part A payments</i>	p = 0.794	p = 0.902	p = 0.769	p = 0.974	p = 0.948	p = 0.778
	<i>Other Part B payments</i>	p = 0.093	p = 0.054	p = 0.065	p = 0.034	p = 0.958	p = 0.735
	<i>30-day PEP payments</i>	p = 0.012	p = 0.019	p = 0.004	p = 0.004	p = 0.603	p = 0.814
Utilization	<i>First PAC SNF</i>	p = 0.165	p = 0.182	p = 0.142	p = 0.27	p = 0.063	p = 0.047
	<i>First PAC IRF</i>	p = 0.050	p = 0.452	p = 0.034	p = 0.627	p = 0.051	p = 0.056
	<i>First PAC HH</i>	p = 0.007	p = 0.014	p = 0.008	p = 0.011	p = 0.597	p = 0.322

Domain	Measure	All LEJR		Elective		Fracture	
		Joint test	Linear test	Joint test	Linear test	Joint test	Linear test
Utilization	<i>First PAC PT/OT</i>	p = 0.476	p = 0.278	p = 0.470	p = 0.277	p = 0.887	p = 0.810
	<i>First PAC home without HH</i>	p = 0.609	p = 0.823	p = 0.501	p = 0.891	p = 0.420	p = 0.715
	<i>SNF days</i>	p = 0.129	p = 0.061	p = 0.164	p = 0.106	p = 0.391	p = 0.130
	<i>IRF days</i>	p = 0.334	p = 0.303	p = 0.399	p = 0.382	p = 0.850	p = 0.572
	<i>Outpatient PT/OT visits</i>	p = 0.214	p = 0.094	p = 0.207	p = 0.075	p = 0.232	p = 0.958
	<i>Any HH use</i>	p = 0.077	p = 0.026	p = 0.115	p = 0.040	p = 0.790	p = 0.668
Quality	<i>Unplanned readmission rate</i>	p = 0.463	p = 0.403	p = 0.310	p = 0.339	p = 0.930	p = 0.97
	<i>ED use rate</i>	p = 0.600	p = 0.757	p = 0.530	p = 0.970	p = 0.695	p = 0.364
	<i>Complication rate</i>	p = 0.753	p = 0.227	p = 0.915	p = 0.597	p = 0.173	p = 0.263
	<i>Mortality rate</i>	p = 0.108	p = 0.678	p = 0.658	p = 0.851	p = 0.165	p = 0.662

Notes: The p-values in this exhibit were the result of risk-adjusted regression models analyzing if the respective CJR and control groups followed parallel trends during the baseline period. For the joint test, we report the p-value of an F-test that tests if the differential between the CJR and control group were jointly equal across annual time periods. For the linear test, we report the p-value of a linear slope coefficient of the quarterly difference between the CJR and control group. Estimates that were significant at the 1%, 5%, or 10% significance level are indicated by shaded cells, respectively. We considered outcomes to fail parallel trends if we rejected the null hypothesis of seemingly parallel trends for both tests at the 10% significance level.

CJR = Comprehensive Care for Joint Replacement; ED = emergency department; HH = home health; IRF = inpatient rehabilitation facility; LEJR = lower extremity joint replacement; OP = outpatient; OT = occupational therapy; p = p-value; PAC = post-acute care; PEP = post episode period; PT = physical therapy; SNF = skilled nursing facility.

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).

Hypothetical DiD Method

Exhibit D-2: Level of trust in DiD estimates, PY 6–PY 8, all LEJR episodes

Domain	Outcome	Differential trend				Hypothetical PY 6–PY 8 DiD				Standard PY 6–PY 8 DiD		Level of trust	Likely sign of true effect
		Coefficient	p-value	90% LCI	90% UCI	Impact	p-value	90% LCI	90% UCI	Impact	p-value		
Payment	<i>Total episode payments</i>	-\$11	0.69	-\$56	\$34	-\$582	0.41	-\$1,751	\$588	-\$998	0.12	Very High	
	<i>SNF payments</i>	-\$10	0.56	-\$37	\$17	\$215	0.74	-\$868	\$1,298	-\$96	0.77	Very High	
	<i>IRF payments</i>	\$12	0.64	-\$32	\$56	-\$995	0.23	-\$2,366	\$376	-\$556	0.11	Very High	
	<i>HH payments</i>	\$8	<0.01	\$4	\$12	-\$365	<0.01	-\$539	-\$192	\$166	0.39	Very Low	- or 0
	<i>Readmission payments</i>	-\$6	0.15	-\$13	\$1	\$135	0.43	-\$147	\$417	-\$126	0.21	Very High	
	<i>Anchor payments</i>	-\$3	0.26	-\$8	\$2	\$109	0.39	-\$101	\$319	-\$11	0.71	High	
	<i>Other Part A payments</i>	\$0	1.00	-\$5	\$5	\$60	0.60	-\$128	\$248	\$60	0.16	Very High	
	<i>Other Part B payments</i>	-\$13	0.05	-\$23	-\$2	\$298	0.07	\$25	\$572	-\$232	0.11	Very Low	+ or 0
	<i>30-day PEP payments</i>	-\$8	0.03	-\$14	-\$2	\$374	<0.01	\$136	\$613	\$55	0.05	Very Low	+
Utilization	<i>First PAC SNF</i>	-0.27 pp	0.07	-0.51 pp	-0.02 pp	7.51 pp	0.24	-3.02 pp	18.05 pp	-0.90 pp	0.67	Very Low	+ or 0
	<i>First PAC IRF</i>	0.07 pp	0.73	-0.26 pp	0.40 pp	-5.11 pp	0.41	-15.27 pp	5.06 pp	-3.82 pp	0.11	Very High	
	<i>First PAC HH</i>	0.12 pp	0.15	-0.02 pp	0.26 pp	0.98 pp	0.77	-4.52 pp	6.48 pp	11.41 pp	0.03	Very Low	+ or 0

Domain	Outcome	Differential trend				Hypothetical PY 6–PY 8 DiD				Standard PY 6–PY 8 DiD		Level of trust	Likely sign of true effect
		Coefficient	p-value	90% LCI	90% UCI	Impact	p-value	90% LCI	90% UCI	Impact	p-value		
Utilization	<i>First PAC OP PT/OT</i>	0.02 pp	0.86	-0.13 pp	0.16 pp	-0.74 pp	0.81	-5.84 pp	4.37 pp	-4.41 pp	0.12	Very High	
	<i>First PAC none</i>	0.06 pp	0.47	-0.08 pp	0.20 pp	-2.65 pp	0.18	-5.86 pp	0.56 pp	-2.28 pp	0.24	Very High	
	<i>Any HH use</i>	0.31 pp	0.02	0.08 pp	0.54 pp	-4.68 pp	0.22	-11.01 pp	1.65 pp	8.14 pp	0.19	Very Low	0
	<i>SNF days</i>	0.09	0.06	0.01	0.17	-2.84	0.18	-6.32	0.64	0.49	0.57	Very Low	- or 0
	<i>IRF days</i>	-0.02	0.17	-0.03	0.00	0.59	0.20	-0.17	1.35	-0.01	0.96	Very High	
	<i>OP PT/OT visits</i>	-0.02	0.23	-0.05	0.01	0.89	0.19	-0.22	2.01	0.11	0.82	High	
Quality	<i>ED use</i>	0.01 pp	0.82	-0.05	0.07	-0.40	0.79	-2.83 pp	2.04 pp	0.05 pp	0.90	Very High	
	<i>Readmissions</i>	-0.04 pp	0.30	-0.10 pp	0.02 pp	0.95 pp	0.49	-1.30 pp	3.20 pp	-0.44 pp	0.19	Very High	
	<i>Complications</i>	0.02 pp	0.21	-0.01 pp	0.06 pp	-1.05 pp	0.20	-2.40 pp	0.30 pp	-0.18 pp	0.33	Very High	
	<i>Mortality</i>	0.01 pp	0.73	-0.02 pp	0.04 pp	-0.29 pp	0.66	-1.38 pp	0.79 pp	-0.08 pp	0.46	Very High	

Notes: The “Coefficient” column shows the estimated differential trend between CJR and control in the baseline period, the “P-value” shows the associated p-value, and the “90% CI LCI” and “90% CI UCI” columns show the 90% lower and upper confidence limits for the estimate, respectively. The “DiD Impact” estimates were reproduced from exhibits in [Appendix D: Additional Findings, Section A.2](#) (below). The “Hypothetical DiD” column is calculated using two-staged least squares DiD. Estimates that were significant at the 1%, 5%, or 10% significance level are indicated by shaded cells, respectively. The “Level of Trust in DiD Estimate” column was subjective and based on the other information in the table. The “Likely sign of true effect” column is the most likely sign of the true effect based on the other information in the table (where “0” indicates a null effect, “+” indicates a positive effect, and “-” indicates a negative effect).

CJR = Comprehensive Care for Joint Replacement; DiD = difference-in-differences; ED = emergency department; HH = home health; IRF = inpatient rehabilitation facility; LCI = lower confidence interval; LEJR = lower extremity joint replacement; OP = outpatient; OT = occupational therapy; PAC = post-acute care; PEP = post episode period; pp = percentage point; PT = physical therapy; PY = performance year; SNF = skilled nursing facility; UCI = upper confidence interval.

Source: CJR evaluation team analysis of Medicare claims and enrollment data for episodes initiated between January 1, 2012, and December 31, 2014, that ended between April 1, 2012, and March 31, 2015 (baseline) and episodes that ended between October 1, 2021, and December 31, 2024 (PY 6–PY 8 intervention).

Exhibit D-3: Level of trust in DiD estimates, PY 6–PY 8, elective LEJR episodes

Domain	Outcome	Differential trend				Hypothetical PY 6–PY 8 DiD				Standard PY 6–PY 8 DiD		Level of trust	Likely sign of true effect
		Coefficient	p-value	90% LCI	90% UCI	Impact	p-value	90% LCI	90% UCI	Impact	p-value		
Payment	<i>Total episode payments</i>	-\$27	0.42	-\$82	\$28	-\$179	0.83	-\$1,561	\$1,204	-\$1,226	0.09	Very High	
	<i>SNF payments</i>	-\$11	0.41	-\$33	\$11	-\$25	0.97	-\$1,001	\$951	-\$300	0.29	Very High	
	<i>IRF payments</i>	\$2	0.92	-\$37	\$42	-\$504	0.49	-\$1,718	\$710	-\$443	0.10	High	
	<i>HH payments</i>	\$9	<0.01	\$4	\$14	-\$394	<0.01	-\$592	-\$195	\$172	0.41	Very Low	- or 0
	<i>Readmission payments</i>	-\$9	0.09	-\$18	\$0	\$249	0.20	-\$70	\$568	-\$121	0.21	Low	0
	<i>Anchor payments</i>	-\$4	0.25	-\$9	\$2	\$98	0.49	-\$137	\$333	-\$32	0.39	High	
	<i>Other Part A payments</i>	\$0	0.89	-\$4	\$5	\$20	0.85	-\$160	\$201	\$34	0.14	Very High	
	<i>Other Part B payments</i>	-\$17	0.03	-\$30	-\$4	\$426	0.04	\$83	\$769	-\$280	0.07	Very Low	+ or 0
	<i>30-Day PEP payments</i>	-\$10	<0.01	-\$15	-\$4	\$372	<0.01	\$162	\$581	-\$3	0.92	Very Low	+
Utilization	<i>First PAC SNF</i>	-0.27 pp	0.08	-0.52 pp	-0.01 pp	6.84 pp	0.31	-4.17 pp	17.85 pp	-1.89 pp	0.45	Very Low	0
	<i>First PAC IRF</i>	0.01 pp	0.98	-0.32 pp	0.33 pp	-2.40 pp	0.69	-12.46 pp	7.65 pp	-3.22 pp	0.11	High	
	<i>First PAC HH</i>	0.15 pp	0.15	-0.02 pp	0.33 pp	-0.71 pp	0.85	-6.88 pp	5.45 pp	12.84 pp	0.02	Very High	

Domain	Outcome	Differential trend				Hypothetical PY 6–PY 8 DiD				Standard PY 6–PY 8 DiD		Level of trust	Likely sign of true effect
		Coefficient	p-value	90% LCI	90% UCI	Impact	p-value	90% LCI	90% UCI	Impact	p-value		
Utilization	<i>First PAC OP PT/OT</i>	0.03 pp	0.75	-0.13 pp	0.20 pp	-0.57 pp	0.87	-6.36 pp	5.23 pp	-5.02 pp	0.12	Very High	
	<i>First PAC none</i>	0.08 pp	0.40	-0.08 pp	0.23 pp	-3.16 pp	0.14	-6.71 pp	0.40 pp	-2.71 pp	0.21	Very High	
	<i>Any HH use</i>	0.34 pp	0.03	0.08 pp	0.60 pp	-4.80 pp	0.29	-12.20 pp	2.60 pp	9.48 pp	0.18	Very Low	0
	<i>SNF days</i>	0.06	0.14	-0.01	0.14	-2.73	0.08	-5.26	-0.20	-0.33	0.65	Very High	
	<i>IRF days</i>	-0.02	0.22	-0.04	0.01	0.80	0.20	-0.22	1.82	0.06	0.80	Very High	
	<i>OP PT/OT visits</i>	-0.02	0.21	-0.05	0.01	0.91	0.18	-0.21	2.03	0.09	0.86	Very High	
Quality	<i>ED use</i>	0.00 pp	0.94	-0.07 pp	0.06 pp	0.19 pp	0.90	-2.35 pp	2.73 pp	0.22 pp	0.58	Very High	
	<i>Readmissions</i>	-0.04 pp	0.25	-0.11 pp	0.02 pp	1.24 pp	0.39	-1.11 pp	3.59 pp	-0.32 pp	0.24	Very High	
	<i>Complications</i>	0.01 pp	0.53	-0.02 pp	0.05 pp	-0.59 pp	0.50	-2.01 pp	0.84 pp	-0.15 pp	0.36	Very High	
	<i>Mortality</i>	0.00 pp	0.89	-0.02 pp	0.02 pp	-0.10 pp	0.81	-0.77 pp	0.57 pp	-0.04 pp	0.35	Very High	

Notes: The “Coefficient” column shows the estimated differential trend between CJR and control in the baseline period, the “P-value” shows the associated p-value, and the “90% CI LCI” and “90% CI UCI” columns show the 90% lower and upper confidence limits for the estimate, respectively. The “DiD Impact” estimates were reproduced from exhibits in [Appendix D: Additional Findings, Section D.2](#) (below). The “Hypothetical DiD” column is calculated using two-staged least squares DiD. Estimates that were significant at the 1%, 5%, or 10% significance level are indicated by shaded cells, respectively. The “Level of Trust in DiD Estimate” column was subjective and based on the other information in the table. The “Likely sign of true effect” column is the most likely sign of the true effect based on the other information in the table (where “0” indicates a null effect, “+” indicates a positive effect, and “-” indicates a negative effect).

CJR = Comprehensive Care for Joint Replacement; DiD = difference-in-differences; ED = emergency department; HH = home health; IRF = inpatient rehabilitation facility; LCI = lower confidence interval; LEJR = lower extremity joint replacement; OP = outpatient; OT = occupational therapy; PAC = post-acute care; PEP = post episode period; pp = percentage point; PT = physical therapy; PY = performance year; SNF = skilled nursing facility; UCI = upper confidence interval.

Source: CJR evaluation team analysis of Medicare claims and enrollment data for episodes initiated between January 1, 2012, and December 31, 2014, that ended between April 1, 2012, and March 31, 2015 (baseline) and episodes that ended between October 1, 2021, and December 31, 2024 (PY 6–PY 8 intervention).

Exhibit D-4: Level of trust in DiD estimates, PY 6–PY 8, fracture LEJR episodes

Domain	Outcome	Differential trend				Hypothetical PY 6–PY 8 DiD				Standard PY 6–PY 8 DiD		Level of trust	Likely sign of true effect
		Coefficient	p-value	90% LCI	90% UCI	Impact	p-value	90% LCI	90% UCI	Impact	p-value		
Payment	<i>Total episode payments</i>	\$45	0.28	-\$24	\$114	-\$1,891	0.33	-\$5,084	\$1,302	-\$111	0.89	High	
	<i>SNF payments</i>	-\$48	0.35	-\$134	\$38	\$3,184	0.06	\$364	\$6,003	\$1,341	0.08	High	
	<i>IRF payments</i>	\$70	0.13	-\$7	\$147	-\$4,322	<0.01	-\$6,764	-\$1,881	-\$1,232	0.11	Low	-
	<i>HH payments</i>	\$4	0.44	-\$5	\$13	-\$148	0.53	-\$532	\$236	-\$16	0.82	Very High	
	<i>Readmission payments</i>	\$12	0.46	-\$15	\$40	-\$568	0.42	-\$1,726	\$590	-\$115	0.44	Very High	
	<i>Anchor payments</i>	-\$3	0.69	-\$15	\$9	\$157	0.59	-\$320	\$633	\$46	0.45	Very High	
	<i>Other Part A payments</i>	\$1	0.89	-\$16	\$19	\$121	0.75	-\$511	\$752	\$174	0.24	Very High	
	<i>Other Part B payments</i>	\$6	0.55	-\$10	\$22	-\$195	0.65	-\$909	\$519	\$36	0.82	Very High	
	<i>30-day PEP payments</i>	-\$0	0.97	-\$25	\$24	\$423	0.48	-\$554	\$1,401	\$404	<0.01	Very High	
Utilization	<i>First PAC SNF</i>	-0.38 pp	0.06	-0.72 pp	-0.05 pp	17.08 pp	0.01	5.94 pp	28.23 pp	2.18 pp	0.41	Very Low	+
	<i>First PAC IRF</i>	0.44 pp	0.07	0.05 pp	0.82 pp	-22.35 pp	<0.01	-35.39 pp	-9.31 pp	-5.70 pp	0.11	Very Low	-
	<i>First PAC HH</i>	-0.08 pp	0.30	-0.20 pp	0.05 pp	6.36 pp	0.04	1.20 pp	11.53 pp	3.38 pp	<0.01	Very High	
	<i>First PAC OP PT/OT</i>	0.01 pp	0.71	-0.05 pp	0.07 pp	-0.75 pp	0.60	-3.13 pp	1.63 pp	-0.08 pp	0.72	Very High	

Domain	Outcome	Differential trend				Hypothetical PY 6–PY 8 DiD				Standard PY 6–PY 8 DiD		Level of trust	Likely sign of true effect
		Coefficient	p-value	90% LCI	90% UCI	Impact	p-value	90% LCI	90% UCI	Impact	p-value		
Utilization	<i>First PAC none</i>	0.01 pp	0.84	-0.08 pp	0.10 pp	-0.34 pp	0.86	-3.59 pp	2.91 pp	0.22 pp	0.59	Very High	
	<i>Any HH use</i>	0.11 pp	0.55	-0.19 pp	0.40 pp	-5.17 pp	0.43	-15.83 pp	5.49 pp	-0.96 pp	0.34	Very High	
	<i>SNF days</i>	0.13	0.14	-0.01	0.27	-3.67	0.30	-9.44	2.10	1.59	0.19	Very High	
	<i>IRF days</i>	-0.02	0.56	-0.06	0.03	0.49	0.66	-1.32	2.29	-0.17	0.62	Very High	
	<i>OP PT/OT visits</i>	0.00	0.91	-0.06	0.07	0.14	0.93	-2.31	2.58	0.30	0.43	Very High	
Quality	<i>ED use</i>	0.07 pp	0.45	-0.09 pp	0.23 pp	-3.84 pp	0.35	-10.53 pp	2.85 pp	-0.88 pp	0.31	Very High	
	<i>Readmissions</i>	0.00 pp	0.99	-0.16 pp	0.16 pp	-1.15 pp	0.78	-7.96 pp	5.66 pp	-1.12 pp	0.27	Very High	
	<i>Complications</i>	0.09 pp	0.26	-0.04 pp	0.22 pp	-3.62 pp	0.26	-8.95 pp	1.71 pp	-0.31 pp	0.53	Very High	
	<i>Mortality</i>	0.01 pp	0.86	-0.12 pp	0.15 pp	-0.85 pp	0.79	-6.20 pp	4.49 pp	-0.28 pp	0.63	Very High	

Notes: The “Coefficient” column shows the estimated differential trend between CJR and control in the baseline period, the “P-value” shows the associated p-value, and the “90% CI LCI” and “90% CI UCI” columns show the 90% lower and upper confidence limits for the estimate, respectively. The “DiD Impact” estimates were reproduced from exhibits in [Appendix D: Additional Findings, Section D.2](#) (below). The “Hypothetical DiD” column is calculated using two-staged least squares DiD. Estimates that were significant at the 1%, 5%, or 10% significance level are indicated by shaded cells, respectively. The “Level of Trust in DiD Estimate” column was subjective and based on the other information in the table. The “Likely sign of true effect” column is the most likely sign of the true effect based on the other information in the table (where “0” indicates a null effect, “+” indicates a positive effect, and “-” indicates a negative effect).

CJR = Comprehensive Care for Joint Replacement; DiD = difference-in-differences; ED = emergency department; HH = home health; IRF = inpatient rehabilitation facility; LCI = lower confidence interval; LEJR = lower extremity joint replacement; OP = outpatient; OT = occupational therapy; PAC = post-acute care; PEP = post episode period; pp = percentage point; PT = physical therapy; PY = performance year; SNF = skilled nursing facility; UCI = upper confidence interval.

Source: CJR evaluation team analysis of Medicare claims and enrollment data for episodes initiated between January 1, 2012, and December 31, 2014, that ended between April 1, 2012, and March 31, 2015 (baseline) and episodes that ended between October 1, 2021, and December 31, 2024 (PY 6–PY 8 intervention).

2. Impact Estimates

Exhibit D-5: Cumulative impact estimates, PY 6–PY 8, all LEJR episodes

Outcome	CJR				Control				PY 6–PY 8 impact (DiD)	DiD as a percent	p- value	90% LCI	90% UCI
	Baseline episodes (N)	PY 6–PY 8 episodes (N)	Baseline mean	PY 6–PY 8 mean	Baseline episodes (N)	PY 6–PY 8 episodes (N)	Baseline mean	PY 6–PY 8 mean					
Total episode payments	114,525	143,369	\$28,508	\$25,667	141,473	151,250	\$27,623	\$25,779	-\$998	-3.5%	0.12	-\$2,046	\$51
SNF payments	114,525	143,369	\$5,874	\$2,970	141,473	151,250	\$5,874	\$3,066	-\$96	-1.6%	0.77	-\$622	\$431
IRF payments	114,525	143,369	\$2,194	\$1,386	141,473	151,250	\$1,963	\$1,712	-\$556	-25.4%	0.11	-\$1,136	\$24
HH payments^a	114,525	143,369	\$2,367	\$1,987	141,473	151,250	\$2,147	\$1,602	\$166	7.0%	0.39	-\$149	\$482
Readmission payments	114,525	143,369	\$1,173	\$1,039	141,473	151,250	\$1,048	\$1,041	-\$126	-10.8%	0.21	-\$291	\$38
Anchor payments	114,525	143,369	\$12,103	\$12,969	141,473	151,250	\$12,096	\$12,974	-\$11	-0.1%	0.71	-\$60	\$38
Other Part A payments	114,525	143,369	\$119	\$115	141,473	151,250	\$168	\$104	\$60	50.2%	0.16	-\$10	\$129
Other Part B payments^a	114,525	143,369	\$4,963	\$4,753	141,473	151,250	\$4,757	\$4,779	-\$232	-4.7%	0.11	-\$472	\$7
30-day PEP payments^a	114,525	143,369	\$1,424	\$1,832	141,473	151,250	\$1,448	\$1,801	\$55	3.9%	0.05	\$10	\$101
First PAC SNF^a	113,951	143,004	41.8%	14.3%	140,987	150,801	41.0%	14.4%	-0.90 pp	-2.2%	0.67	-4.39 pp	2.59 pp
First PAC IRF	113,951	143,004	13.6%	5.5%	140,987	150,801	12.2%	7.9%	-3.82 pp	-28.0%	0.11	-7.73 pp	0.09 pp
First PAC HH^a	113,951	143,004	36.2%	54.6%	140,987	150,801	33.5%	40.5%	11.41 pp	31.5%	0.03	2.95 pp	19.88 pp
First PAC OP PT/OT	113,951	143,004	5.7%	19.0%	140,987	150,801	9.1%	26.8%	-4.41 pp	-78.1%	0.12	-9.04 pp	0.22 pp
First PAC none	113,951	143,004	2.7%	6.7%	140,987	150,801	4.2%	10.5%	-2.28 pp	-83.7%	0.24	-5.49 pp	0.93 pp

Outcome	CJR				Control				PY 6–PY 8 impact (DiD)	DiD as a percent	p- value	90% LCI	90% UCI
	Baseline episodes (N)	PY 6–PY 8 episodes (N)	Baseline mean	PY 6–PY 8 mean	Baseline episodes (N)	PY 6–PY 8 episodes (N)	Baseline mean	PY 6–PY 8 mean					
Any HH use ^a	114,525	143,369	73.6%	70.7%	141,473	151,250	68.3%	57.2%	8.14 pp	11.1%	0.19	-2.11 pp	18.38 pp
SNF days ^a	51,542	21,273	27.6	24.7	58,508	19,293	28.4	24.9	0.49	1.8%	0.57	-0.94	1.93
IRF days	16,415	8,445	11.7	12.2	17,844	11,436	11.4	11.9	-0.01	-0.1%	0.96	-0.51	0.49
OP PT/OT visits	69,289	122,591	13.0	14.7	90,157	130,926	13.3	15.0	0.11	0.9%	0.82	-0.71	0.93
ED use rate	114,500	143,338	12.7%	13.1%	141,450	151,228	12.2%	12.4%	0.05 pp	0.4%	0.90	-0.59 pp	0.69 pp
Unplanned readmission rate	114,500	143,338	9.8%	7.0%	141,450	151,228	9.5%	7.0%	-0.44 pp	-4.5%	0.19	-1.00 pp	0.11 pp
Complication rate	114,500	143,338	3.9%	2.3%	141,450	151,228	3.7%	2.2%	-0.18 pp	-4.5%	0.33	-0.47 pp	0.12 pp
Mortality rate	117,415	145,637	2.3%	2.1%	144,680	153,559	2.4%	2.2%	-0.08 pp	-3.4%	0.46	-0.26 pp	0.10 pp

Notes: The estimates in this exhibit were the result of a DiD model. DiD Estimates that were significant at the 1%, 5%, or 10% significance level are indicated by shaded cells, respectively. Outcome definitions are presented in [Appendix B: Data and Methods](#).

CJR = Comprehensive Care for Joint Replacement; DiD = difference-in-differences; HH = home health; IRF = inpatient rehabilitation facility; LCI = lower confidence interval; LEJR = lower extremity joint replacement; N = number; PY = performance year; SNF = skilled nursing facility; UCI = upper confidence interval.

^a Indicates that after a holistic approach towards analyzing parallel trends, we believe the CJR and control populations may have been on relatively large differential trends in the baseline period for this outcome. This could lead to this estimate not being an unbiased causal estimate of the CJR Model.

^b Percents are calculated by dividing the impact by the risk-adjusted mean value of the outcome for CJR episodes in the baseline period.

Source: CJR evaluation team analysis of Medicare claims and enrollment data for episodes initiated between January 1, 2012, and December 31, 2014, that ended between April 1, 2012, and March 31, 2015 (baseline) and episodes that ended between October 1, 2021, and December 31, 2024 (PY 6–PY 8 intervention).

Exhibit D-6: Cumulative impact estimates, PY 6–PY 8, elective LEJR episodes

Outcome	CJR				Control				PY 6–PY 8 impact (DiD)	DiD as a percent	p- value	90% LCI	90% UCI
	Baseline episodes (N)	PY 6–PY 8 episodes (N)	Baseline mean	PY 6–PY 8 mean	Baseline episodes (N)	PY 6–PY 8 episodes (N)	Baseline mean	PY 6–PY 8 mean					
Total episode payments	96,266	126,930	\$25,714	\$21,881	122,534	135,008	\$24,578	\$21,971	-\$1,226	-4.8%	0.09	-\$2,422	-\$30
SNF payments	96,266	126,930	\$4,139	\$1,280	122,534	135,008	\$3,969	\$1,411	-\$300	-7.2%	0.29	-\$768	\$168
IRF payments	96,266	126,930	\$1,657	\$686	122,534	135,008	\$1,452	\$924	-\$443	-26.7%	0.10	-\$885	-\$1
HH payments^a	96,266	126,930	\$2,364	\$1,908	122,534	135,008	\$2,124	\$1,497	\$172	7.3%	0.41	-\$175	\$518
Readmission payments^a	96,266	126,930	\$947	\$780	122,534	135,008	\$831	\$786	-\$121	-12.8%	0.21	-\$280	\$37
Anchor payments	96,266	126,930	\$11,939	\$12,640	122,534	135,008	\$11,909	\$12,641	-\$32	-0.3%	0.39	-\$93	\$29
Other Part A payments	96,266	126,930	\$56	\$36	122,534	135,008	\$82	\$27	\$34	61.4%	0.14	-\$4	\$73
Other Part B payments^a	96,266	126,930	\$4,779	\$4,404	122,534	135,008	\$4,539	\$4,445	-\$280	-5.9%	0.07	-\$534	-\$27
30-day PEP payments^a	96,266	126,922	\$1,125	\$1,499	122,534	135,006	\$1,120	\$1,497	-\$3	-0.2%	0.92	-\$44	\$39
First PAC SNF^a	95,879	126,676	38.4%	7.8%	122,289	134,698	36.8%	8.1%	-1.89 pp	-4.9%	0.45	-6.00 pp	2.21 pp
First PAC IRF	95,879	126,676	11.5%	2.5%	122,289	134,698	10.3%	4.6%	-3.22 pp	-27.9%	0.11	-6.54 pp	0.09 pp
First PAC HH	95,879	126,676	40.8%	60.7%	122,289	134,698	37.9%	45.0%	12.84 pp	31.5%	0.02	3.55 pp	22.13 pp
First PAC OP PT/OT	95,879	126,676	6.3%	21.6%	122,289	134,698	10.4%	30.6%	-5.02 pp	-79.1%	0.12	-10.34 pp	0.31 pp
First PAC none	95,879	126,676	2.9%	7.3%	122,289	134,698	4.6%	11.7%	-2.71 pp	-94.1%	0.21	-6.29 pp	0.87 pp
Any HH use^a	96,266	126,930	74.2%	70.9%	122,534	135,008	68.6%	55.8%	9.48 pp	12.8%	0.18	-2.20 pp	21.15 pp
SNF days	38,172	11,191	20.4	19.4	45,034	10,171	20.7	20.0	-0.33	-1.6%	0.65	-1.56	0.89

Outcome	CJR				Control				PY 6–PY 8 impact (DiD)	DiD as a percent	p-value	90% LCI	90% UCI
	Baseline episodes (N)	PY 6–PY 8 episodes (N)	Baseline mean	PY 6–PY 8 mean	Baseline episodes (N)	PY 6–PY 8 episodes (N)	Baseline mean	PY 6–PY 8 mean					
IRF days	11,325	4,138	10.2	11.4	12,356	6,153	10.1	11.2	0.06	0.6%	0.80	-0.36	0.48
OP PT/OT visits	64,933	116,759	13.1	14.9	85,782	125,350	13.5	15.1	0.09	0.7%	0.86	-0.74	0.93
ED use rate	96,244	126,913	11.9%	12.2%	122,516	134,992	11.4%	11.5%	0.22 pp	1.8%	0.58	-0.43 pp	0.86 pp
Unplanned readmission rate	96,244	126,913	8.1%	5.5%	122,516	134,992	7.8%	5.6%	-0.32 pp	-3.9%	0.24	-0.77 pp	0.13 pp
Complication rate	96,244	126,905	3.0%	1.7%	122,516	134,990	2.7%	1.6%	-0.15 pp	-5.1%	0.36	-0.42 pp	0.12 pp
Mortality rate	96,783	127,257	0.5%	0.4%	123,165	135,408	0.5%	0.4%	-0.04 pp	-8.4%	0.35	-0.12 pp	0.03 pp

Notes: The estimates in this exhibit were the result of a DiD model. DiD Estimates that were significant at the 1%, 5%, or 10% significance level are indicated by shaded cells, respectively. Outcome definitions are presented in [Appendix B: Data and Methods](#).

CJR = Comprehensive Care for Joint Replacement; DiD = difference-in-differences; ED = emergency department; HH = home health; IRF = inpatient rehabilitation facility; LCI = lower confidence interval; LEJR = lower extremity joint replacement; N = number; OP = outpatient; PEP = post-episode period; pp = percentage point; PT = physical therapy; PY = performance year; SNF = skilled nursing facility; UCI = upper confidence interval.

^a Indicates that after a holistic approach towards analyzing parallel trends, we believe the CJR and control populations may have been on relatively large differential trends in the baseline period for this outcome. This could lead to this estimate not being an unbiased causal estimate of the CJR Model.

^b Percents are calculated by dividing the impact by the risk-adjusted mean value of the outcome for CJR episodes in the baseline period.

Source: CJR evaluation team analysis of Medicare claims and enrollment data for episodes initiated between January 1, 2012, and December 31, 2014, that ended between April 1, 2012, and March 31, 2015 (baseline) and episodes that ended between October 1, 2021, and December 31, 2024 (PY 6–PY 8 intervention).

Exhibit D-7: Cumulative impact estimates, PY 6–PY 8, fracture LEJR episodes

Outcome	CJR				Control				PY 6–PY 8 impact (DiD)	DiD as a percent	p-value	90% LCI	90% UCI
	Baseline episodes (N)	PY 6–PY 8 episodes (N)	Baseline mean	PY 6–PY 8 mean	Baseline episodes (N)	PY 6–PY 8 episodes (N)	Baseline mean	PY 6–PY 8 mean					
Total episode payments	18,259	16,350	\$47,285	\$50,556	18,939	16,104	\$47,280	\$50,663	-\$111	-0.4%	0.89	-\$1,442	\$1,219
SNF payments	18,259	16,350	\$16,630	\$15,717	18,939	16,104	\$17,621	\$15,368	\$1,341	8.1%	0.08	\$93	\$2,589
IRF payments^a	18,259	16,350	\$5,444	\$6,577	18,939	16,104	\$4,893	\$7,258	-\$1,232	-22.6%	0.11	-\$2,492	\$27
HH payments	18,259	16,350	\$2,448	\$2,501	18,939	16,104	\$2,349	\$2,418	-\$16	-0.7%	0.82	-\$135	\$102
Readmission payments	18,259	16,350	\$2,602	\$2,794	18,939	16,104	\$2,459	\$2,766	-\$115	-4.4%	0.44	-\$361	\$131
Anchor payments	18,259	16,350	\$13,287	\$15,170	18,939	16,104	\$13,332	\$15,169	\$46	0.3%	0.45	-\$56	\$149
Other Part A payments	18,259	16,350	\$557	\$619	18,939	16,104	\$702	\$589	\$174	31.3%	0.24	-\$70	\$418
Other Part B payments	18,259	16,350	\$6,298	\$7,062	18,939	16,104	\$6,178	\$6,906	\$36	0.6%	0.82	-\$222	\$295
30-day PEP payments	18,259	16,341	\$3,396	\$3,944	18,939	16,092	\$3,586	\$3,730	\$404	11.9%	<0.01	\$169	\$639
First PAC SNF^a	18,072	16,239	64.7%	57.0%	18,698	15,966	66.0%	56.0%	2.18 pp	3.4%	0.41	-2.19 pp	6.56 pp
First PAC IRF^a	18,072	16,239	27.7%	24.6%	18,698	15,966	25.9%	28.5%	-5.70 pp	-20.6%	0.11	-11.63 pp	0.23 pp
First PAC HH	18,072	16,239	5.4%	14.4%	18,698	15,966	5.9%	11.6%	3.38 pp	63.1%	<0.01	1.32 pp	5.45 pp
First PAC OP PT/OT	18,072	16,239	0.9%	1.6%	18,698	15,966	0.8%	1.6%	-0.08 pp	-9.1%	0.72	-0.46 pp	0.30 pp
First PAC none	18,072	16,239	1.4%	2.4%	18,698	15,966	1.4%	2.3%	0.22 pp	15.8%	0.59	-0.43 pp	0.87 pp
Any HH use	18,259	16,350	69.8%	69.6%	18,939	16,104	68.0%	68.8%	0.34 pp	-1.4%	0.34	-2.62 pp	-0.70 pp
SNF days	13,370	10,072	43.1	37.4	13,474	9,112	45.1	37.8	1.59	3.7%	0.19	-0.39	3.57

Outcome	CJR				Control				PY 6–PY 8 impact (DiD)	DiD as a percent	p-value	90% LCI	90% UCI
	Baseline episodes (N)	PY 6–PY 8 episodes (N)	Baseline mean	PY 6–PY 8 mean	Baseline episodes (N)	PY 6–PY 8 episodes (N)	Baseline mean	PY 6–PY 8 mean					
IRF days	5,090	4,301	14.1	13.8	5,488	5,265	13.7	13.5	-0.17	-1.2%	0.62	-0.73	0.39
OP PT/OT visits	4,356	5,746	10.5	12.6	4,375	5,452	10.7	12.4	0.30	2.9%	0.43	-0.34	0.95
ED use rate	18,256	16,336	17.9%	18.7%	18,934	16,098	17.2%	18.9%	-0.88 pp	-4.9%	0.31	-2.30 pp	0.55 pp
Unplanned readmission rate	18,256	16,336	21.2%	16.4%	18,934	16,098	20.4%	16.7%	-1.12 pp	-5.3%	0.27	-2.80 pp	0.56 pp
Complication rate	18,256	16,336	10.1%	6.5%	18,934	16,098	9.8%	6.5%	-0.31 pp	-3.1%	0.53	-1.13 pp	0.51 pp
Mortality rate	20,632	18,214	12.7%	11.9%	21,515	17,955	13.1%	12.6%	-0.28 pp	-2.2%	0.63	-1.22 pp	0.66 pp

Notes: The estimates in this exhibit were the result of a DiD model. DiD Estimates that were significant at the 1%, 5%, or 10% significance level are indicated by shaded cells, respectively. Outcome definitions are presented in [Appendix B: Data and Methods](#).

CJR = Comprehensive Care for Joint Replacement; DiD = difference-in-differences; ED = emergency department; HH = home health; IRF = inpatient rehabilitation facility; LCI = lower confidence interval; LEJR = lower extremity joint replacement; N = number; OP = outpatient; PEP = post-episode period; pp = percentage point; PT = physical therapy; PY = performance year; SNF = skilled nursing facility; UCI = upper confidence interval.

^a Indicates that after a holistic approach towards analyzing parallel trends, we believe the CJR and control populations may have been on relatively large differential trends in the baseline period for this outcome. This could lead to this estimate not being an unbiased causal estimate of the CJR Model.

^b Percents are calculated by dividing the impact by the risk-adjusted mean value of the outcome for CJR episodes in the baseline period.

Source: CJR evaluation team analysis of Medicare claims and enrollment data for episodes initiated between January 1, 2012, and December 31, 2014, that ended between April 1, 2012, and March 31, 2015 (baseline) and episodes that ended between October 1, 2021, and December 31, 2024 (PY 6–PY 8 intervention).

Exhibit D-8: Performance year specific impact estimates, PY 6–PY 8, all LEJR episodes

Outcome	PY 6				PY 7				PY 8			
	Impact (DiD)	p-value	90% LCI	90% UCI	Impact (DiD)	p-value	90% LCI	90% UCI	Impact (DiD)	p-value	90% LCI	90% UCI
Total episode payments	-\$984	0.09	-\$1,939	-\$28	-\$922	0.15	-\$1,990	\$146	-\$1,051	0.13	-\$2,205	\$104
SNF payments	-\$129	0.66	-\$617	\$358	-\$79	0.81	-\$613	\$454	-\$77	0.83	-\$658	\$504
IRF payments	-\$571	0.07	-\$1,090	-\$52	-\$475	0.19	-\$1,073	\$123	-\$599	0.13	-\$1,257	\$58
HH payments^a	\$191	0.32	-\$126	\$508	\$161	0.41	-\$162	\$483	\$138	0.47	-\$175	\$452
Readmission payments	-\$166	0.06	-\$310	-\$21	-\$105	0.35	-\$290	\$80	-\$99	0.38	-\$284	\$86
Anchor payments	\$13	0.66	-\$37	\$64	-\$18	0.56	-\$69	\$33	-\$24	0.48	-\$80	\$32
Other Part A payments	\$43	0.25	-\$19	\$106	\$55	0.21	-\$18	\$128	\$83	0.08	\$6	\$161
Other Part B payments^a	-\$220	0.09	-\$434	-\$6	-\$203	0.15	-\$434	\$28	-\$260	0.15	-\$556	\$36
30-day PEP payments^a	\$72	0.03	\$18	\$126	\$55	0.12	-\$3	\$113	\$45	0.21	-\$14	\$104
First PAC SNF^a	-1.05 pp	0.62	-4.50 pp	2.41 pp	-1.23 pp	0.56	-4.69 pp	2.24 pp	-0.41 pp	0.85	-3.98 pp	3.15 pp
First PAC IRF	-3.86 pp	0.08	-7.52 pp	-0.20 pp	-3.43 pp	0.16	-7.44 pp	0.58 pp	-4.14 pp	0.10	-8.32 pp	0.04 pp
First PAC HH^a	12.71 pp	0.01	4.29 pp	21.12 pp	10.91 pp	0.04	2.28 pp	19.53 pp	10.22 pp	0.05	1.59 pp	18.86 pp
First PAC OP PT/OT	-4.39 pp	0.11	-8.90 pp	0.13 pp	-4.40 pp	0.13	-9.13 pp	0.34 pp	-4.26 pp	0.17	-9.34 pp	0.82 pp
First PAC none	-3.41 pp	0.07	-6.45 pp	-0.37 pp	-1.85 pp	0.38	-5.34 pp	1.64 pp	-1.41 pp	0.49	-4.76 pp	1.94 pp
Any HH use^a	8.69 pp	0.16	-1.37 pp	18.75 pp	7.84 pp	0.22	-2.71 pp	18.39 pp	7.66 pp	0.23	-2.76 pp	18.08 pp
SNF days^a	0.33	0.67	-0.95	1.61	0.69	0.44	-0.81	2.20	0.52	0.63	-1.26	2.30
IRF days	0.21	0.49	-0.29	0.70	-0.21	0.44	-0.67	0.24	-0.05	0.90	-0.70	0.60

Outcome	PY 6				PY 7				PY 8			
	Impact (DiD)	p-value	90% LCI	90% UCI	Impact (DiD)	p-value	90% LCI	90% UCI	Impact (DiD)	p-value	90% LCI	90% UCI
OP PT/OT visits	0.01	0.98	-0.80	0.82	0.23	0.63	-0.58	1.05	0.13	0.80	-0.73	0.98
ED use	0.29 pp	0.48	-0.38 pp	0.96 pp	0.04 pp	0.93	-0.78 pp	0.86 pp	-0.21 pp	0.60	-0.89 pp	0.46 pp
Readmissions	-0.49 pp	0.12	-1.00 pp	0.02 pp	-0.32 pp	0.43	-0.99 pp	0.35 pp	-0.51 pp	0.18	-1.13 pp	0.12 pp
Complications	-0.22 pp	0.21	-0.50 pp	0.07 pp	-0.22 pp	0.32	-0.57 pp	0.14 pp	-0.09 pp	0.66	-0.43 pp	0.25 pp
Mortality	-0.03 pp	0.80	-0.23 pp	0.17 pp	-0.08 pp	0.55	-0.31 pp	0.14 pp	-0.15 pp	0.25	-0.38 pp	0.07 pp

Notes: The estimates in this exhibit were the result of a DiD model. DiD Estimates that were significant at the 1%, 5%, or 10% significance level are indicated by shaded cells, respectively. Outcome definitions are presented in [Appendix B: Data and Methods](#).

CJR = Comprehensive Care for Joint Replacement; DiD = difference-in-differences; ED = emergency department; HH = home health; IRF = inpatient rehabilitation facility; LCI = lower confidence interval; LEJR = lower extremity joint replacement; OP = outpatient; OT = occupational therapy; PAC = post-acute care; PEP = post-episode period; pp = percentage point; PT = physical therapy; PY = performance year; SNF = skilled nursing facility; UCI = upper confidence interval.

^a Indicates that after a holistic approach towards analyzing parallel trends, we believe the CJR and control populations may have been on relatively large differential trends in the baseline period for this outcome. This could lead to this estimate not being an unbiased causal estimate of the CJR Model.

^b Percents are calculated by dividing the impact by the risk-adjusted mean value of the outcome for CJR episodes in the baseline period.

Source: CJR evaluation team analysis of Medicare claims and enrollment data for episodes initiated between January 1, 2012, and December 31, 2014, that ended between April 1, 2012, and March 31, 2015 (baseline) and episodes that ended between October 1, 2021, and December 31, 2024 (PY 6–PY 8 intervention).

Exhibit D-9: Performance year specific impact estimates, PY 6–PY 8, elective LEJR episodes

Outcome	PY 6				PY 7				PY 8			
	Impact (DiD)	p-value	90% LCI	90% UCI	Impact (DiD)	p-value	90% LCI	90% UCI	Impact (DiD)	p-value	90% LCI	90% UCI
Total episode payments	-\$1,133	0.08	-\$2,210	-\$55	-\$1,191	0.11	-\$2,432	\$50	-\$1,347	0.09	-\$2,651	-\$43
SNF payments	-\$331	0.21	-\$769	\$107	-\$307	0.29	-\$788	\$174	-\$266	0.39	-\$770	\$239
IRF payments	-\$419	0.08	-\$814	-\$23	-\$410	0.14	-\$867	\$47	-\$489	0.11	-\$989	\$12
HH payments^a	\$206	0.33	-\$143	\$554	\$158	0.46	-\$194	\$510	\$140	0.50	-\$203	\$484
Readmission payments^a	-\$149	0.08	-\$290	-\$8	-\$113	0.29	-\$289	\$63	-\$97	0.36	-\$271	\$77
Anchor payments	-\$8	0.81	-\$65	\$48	-\$35	0.36	-\$98	\$28	-\$50	0.28	-\$127	\$27
Other Part A payments	\$24	0.30	-\$14	\$61	\$40	0.09	\$1	\$78	\$41	0.13	-\$3	\$86
Other Part B payments^a	-\$251	0.07	-\$482	-\$21	-\$247	0.10	-\$493	-\$1	-\$334	0.07	-\$632	-\$36
30-day PEP payments^a	\$20	0.51	-\$29	\$68	-\$20	0.55	-\$76	\$36	-\$4	0.90	-\$52	\$44
First PAC SNF^a	-2.11 pp	0.39	-6.12 pp	1.89 pp	-2.01 pp	0.43	-6.21 pp	2.18 pp	-1.47 pp	0.55	-5.57 pp	2.62 pp
First PAC IRF	-3.17 pp	0.10	-6.33 pp	-0.02 pp	-3.01 pp	0.14	-6.37 pp	0.34 pp	-3.48 pp	0.10	-7.00 pp	0.03 pp
First PAC HH	14.29 pp	0.01	5.03 pp	23.56 pp	12.23 pp	0.03	2.74 pp	21.72 pp	11.52 pp	0.04	2.09 pp	20.94 pp
First PAC OP PT/OT	-5.02 pp	0.11	-10.21 pp	0.18 pp	-5.01 pp	0.13	-10.44 pp	0.42 pp	-4.81 pp	0.18	-10.64 pp	1.03 pp
First PAC none	-3.99 pp	0.05	-7.38 pp	-0.60 pp	-2.19 pp	0.36	-6.11 pp	1.72 pp	-1.75 pp	0.44	-5.47 pp	1.97 pp
Any HH use^a	10.21 pp	0.14	-1.28 pp	21.70 pp	9.08 pp	0.21	-2.87 pp	21.03 pp	8.89 pp	0.22	-2.98 pp	20.77 pp
SNF days	-0.50	0.51	-1.76	0.76	-0.39	0.62	-1.69	0.91	-0.02	0.98	-1.47	1.43
IRF days	0.43	0.11	-0.01	0.87	-0.29	0.25	-0.70	0.12	0.05	0.87	-0.50	0.61

Outcome	PY 6				PY 7				PY 8			
	Impact (DiD)	p-value	90% LCI	90% UCI	Impact (DiD)	p-value	90% LCI	90% UCI	Impact (DiD)	p-value	90% LCI	90% UCI
OP PT/OT visits	-0.01	0.98	-0.83	0.80	0.21	0.68	-0.63	1.04	0.12	0.83	-0.76	0.99
ED use	0.47 pp	0.29	-0.26 pp	1.20 pp	0.17 pp	0.72	-0.62 pp	0.96 pp	-0.01 pp	0.98	-0.64 pp	0.62 pp
Readmissions	-0.27 pp	0.31	-0.71 pp	0.16 pp	-0.27 pp	0.43	-0.85 pp	0.30 pp	-0.41 pp	0.16	-0.89 pp	0.07 pp
Complications	-0.22 pp	0.14	-0.48 pp	0.03 pp	-0.16 pp	0.43	-0.49 pp	0.17 pp	-0.06 pp	0.72	-0.36 pp	0.23 pp
Mortality	-0.02 pp	0.76	-0.11 pp	0.07 pp	-0.07 pp	0.19	-0.16 pp	0.02 pp	-0.04 pp	0.48	-0.15 pp	0.06 pp

Notes: The estimates in this exhibit were the result of a DiD model. DiD Estimates that were significant at the 1%, 5%, or 10% significance level are indicated by shaded cells, respectively. Outcome definitions are presented in [Appendix B: Data and Methods](#).

CJR = Comprehensive Care for Joint Replacement; DiD = difference-in-differences; ED = emergency department; HH = home health; IRF = inpatient rehabilitation facility; LCI = lower confidence interval; LEJR = lower extremity joint replacement; OP = outpatient; OT = occupational therapy; PAC = post-acute care; PEP = post-episode period; pp = percentage point; PT = physical therapy; PY = performance year; SNF = skilled nursing facility; UCI = upper confidence interval.

^a Indicates that after a holistic approach towards analyzing parallel trends, we believe the CJR and control populations may have been on relatively large differential trends in the baseline period for this outcome. This could lead to this estimate not being an unbiased causal estimate of the CJR Model.

^b Percents are calculated by dividing the impact by the risk-adjusted mean value of the outcome for CJR episodes in the baseline period.

Source: CJR evaluation team analysis of Medicare claims and enrollment data for episodes initiated between January 1, 2012, and December 31, 2014, that ended between April 1, 2012, and March 31, 2015 (baseline) and episodes that ended between October 1, 2021, and December 31, 2024 (PY 6–PY 8 intervention).

Exhibit D-10: Performance year specific impact estimates, PY 6–PY 8, fracture LEJR episodes

Outcome	PY 6				PY 7				PY 8			
	Impact (DiD)	p-value	90% LCI	90% UCI	Impact (DiD)	p-value	90% LCI	90% UCI	Impact (DiD)	p-value	90% LCI	90% UCI
Total episode payments	-\$605	0.43	-\$1,864	\$653	\$364	0.70	-\$1,223	\$1,952	\$185	0.84	-\$1,336	\$1,706
SNF payments	\$1,519	0.02	\$471	\$2,567	\$1,338	0.13	-\$119	\$2,796	\$1,181	0.17	-\$245	\$2,606
IRF payments^a	-\$1,540	0.03	-\$2,702	-\$378	-\$791	0.32	-\$2,106	\$524	-\$1,252	0.16	-\$2,723	\$220
HH payments	-\$54	0.47	-\$178	\$69	\$26	0.74	-\$102	\$155	-\$19	0.81	-\$151	\$113
Readmission payments	-\$230	0.14	-\$484	\$24	-\$2	0.99	-\$355	\$352	-\$83	0.69	-\$433	\$266
Anchor payments	\$98	0.14	-\$11	\$207	\$5	0.94	-\$102	\$111	\$51	0.53	-\$83	\$186
Other Part A payments	\$113	0.35	-\$88	\$314	\$87	0.62	-\$202	\$376	\$332	0.07	\$32	\$632
Other Part B payments	-\$73	0.59	-\$300	\$154	\$52	0.77	-\$239	\$343	\$187	0.58	-\$379	\$752
30-day PEP payments	\$368	0.03	\$88	\$647	\$539	<0.01	\$265	\$813	\$336	0.09	\$14	\$658
First PAC SNF^a	3.08 pp	0.20	-0.86 pp	7.03 pp	0.48 pp	0.88	-4.81 pp	5.77 pp	2.72 pp	0.34	-1.93 pp	7.37 pp
First PAC IRF^a	-6.63 pp	0.05	-12.18 pp	-1.08 pp	-4.22 pp	0.28	-10.59 pp	2.14 pp	-5.98 pp	0.12	-12.37 pp	0.41 pp
First PAC HH	3.52 pp	0.01	1.27 pp	5.78 pp	3.58 pp	0.01	1.27 pp	5.89 pp	2.98 pp	0.03	0.72 pp	5.23 pp
First PAC OP PT/OT	-0.11 pp	0.66	-0.50 pp	0.29 pp	0.18 pp	0.57	-0.33 pp	0.68 pp	-0.29 pp	0.36	-0.81 pp	0.23 pp
First PAC none	0.13 pp	0.76	-0.55 pp	0.81 pp	-0.02 pp	0.97	-0.75 pp	0.72 pp	0.57 pp	0.28	-0.30 pp	1.44 pp
Any HH use	-1.37 pp	0.13	-2.83 pp	0.10 pp	-1.04 pp	0.45	-3.33 pp	1.25 pp	-0.55 pp	0.66	-2.60 pp	1.50 pp
SNF days	1.66	0.08	0.08	3.24	1.99	0.18	-0.46	4.43	1.16	0.41	-1.15	3.46
IRF days	-0.12	0.72	-0.69	0.45	-0.19	0.58	-0.77	0.38	-0.21	0.65	-0.97	0.56
OP PT/OT visits	0.23	0.64	-0.60	1.06	0.62	0.11	-0.03	1.27	0.11	0.79	-0.55	0.76
ED use	-0.62 pp	0.49	-2.10 pp	0.87 pp	-0.72 pp	0.48	-2.42 pp	0.97 pp	-1.30 pp	0.30	-3.36 pp	0.76 pp
Readmissions	-1.75 pp	0.08	-3.40 pp	-0.09 pp	-0.53 pp	0.62	-2.26 pp	1.21 pp	-0.91 pp	0.50	-3.15 pp	1.33 pp
Complications	-0.04 pp	0.95	-0.99 pp	0.91 pp	-0.63 pp	0.31	-1.65 pp	0.39 pp	-0.30 pp	0.65	-1.40 pp	0.79 pp
Mortality	0.00 pp	1.00	-1.06 pp	1.06 pp	-0.10 pp	0.90	-1.38 pp	1.18 pp	-0.86 pp	0.27	-2.14 pp	0.42 pp

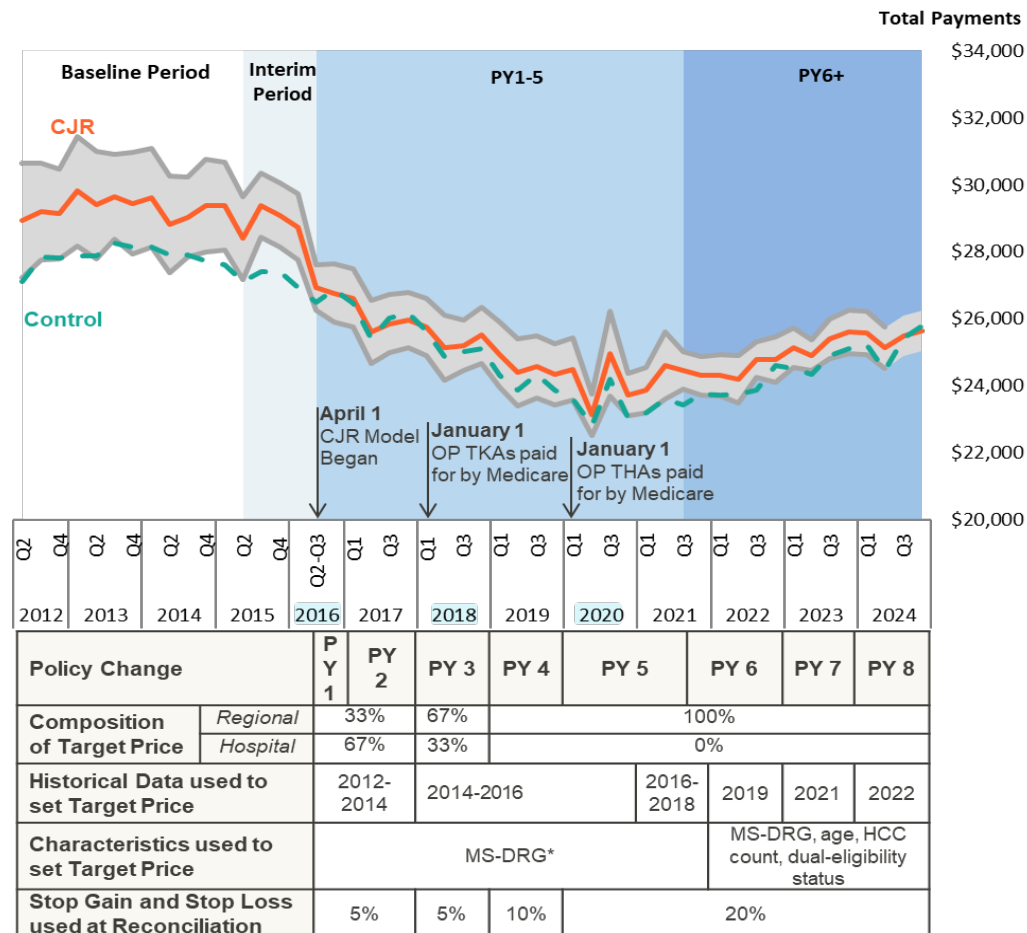
Notes: The estimates in this exhibit were the result of a DiD model. DiD Estimates that were significant at the 1%, 5%, or 10% significance level are indicated by shaded cells, respectively. Outcome definitions are presented in [Appendix B: Data and Methods](#).

CJR = Comprehensive Care for Joint Replacement; DiD = difference-in-differences; ED = emergency department; HH = home health; IRF = inpatient rehabilitation facility; LCI = lower confidence interval; LEJR = lower extremity joint replacement; OP = outpatient; OT = occupational therapy; PAC = post-acute care; PEP = post-episode period; pp = percentage point; PT = physical therapy; PY = performance year; SNF = skilled nursing facility; UCI = upper confidence interval.

^a Indicates that after a holistic approach towards analyzing parallel trends, we believe the CJR and control populations may have been on relatively large differential trends in the baseline period for this outcome. This could lead to this estimate not being an unbiased causal estimate of the CJR Model.

^b Percents are calculated by dividing the impact by the risk-adjusted mean value of the outcome for CJR episodes in the baseline period.

Source: CJR evaluation team analysis of Medicare claims and enrollment data for episodes initiated between January 1, 2012, and December 31, 2014, that ended between April 1, 2012, and March 31, 2015 (baseline) and episodes that ended between October 1, 2021, and December 31, 2024 (PY 6–PY 8 intervention).

Exhibit D-11: Unadjusted total episode payment trends, baseline through PY 8, all LEJR episodes

Notes: The trends in this exhibit are not risk-adjusted, but do use our standard sampling weights. 90% confidence intervals are shown for the control trend line. Both the CJR and control line include values for OP LEJRs since removed from the inpatient only list, even though OP LEJRs were not CJR episodes until the start of PY 6. See [Appendix B: Data and Methods](#) for more information.

CJR = Comprehensive Care for Joint Replacement; HCC = Hierarchical Condition Category; LEJR = lower extremity joint replacement; MS-DRG = Medicare Severity Diagnosis-Related Group; OP = Outpatient; PY = performance year; Q = quarter; THA = total hip arthroplasty; TKA = total knee arthroplasty.

* Prior to elective and fracture LEJRs each getting their own set of MS-DRGs, CJR target prices were also adjusted for fracture status.

Source: CJR evaluation team analysis of Medicare claims and enrollment data for episodes initiated after January 1, 2012, that ended before December 31, 2024.

Exhibit D-12: Episode payment outcomes, PY 6–PY 8, for elective patients who were dually eligible

Measure	Population	CJR				Control				PY 6–PY 8 impact (DiD)	p-value	90% LCI	90% UCI
		Baseline episodes (N)	PY 6–PY 8 episodes (N)	Baseline mean	PY 6–PY 8 mean	Baseline episodes (N)	PY 6–PY 8 episodes (N)	Baseline mean	PY 6–PY 8 mean				
Total episode payments	<i>Dually eligible</i>	12,038	7,721	\$30,614	\$26,596	12,017	6,267	\$28,384	\$25,878	-\$1,513	0.12	-\$3,112	\$87
	<i>Not dually eligible</i>	84,228	119,209	\$25,289	\$21,382	110,517	128,741	\$24,270	\$21,505	-\$1,142	0.11	-\$2,316	\$32
HH payments	<i>Dually eligible</i>	12,038	7,721	\$2,606	\$2,275	12,017	6,267	\$2,533	\$1,911	\$291	0.07	\$28	\$555
	<i>Not dually eligible</i>	84,228	119,209	\$2,353	\$1,870	110,517	128,741	\$2,102	\$1,453	\$166	0.45	-\$194	\$527
SNF payments	<i>Dually eligible</i>	12,038	7,721	\$7,770	\$4,114	12,017	6,267	\$6,399	\$3,604	-\$861	0.09	-\$1,706	-\$17
	<i>Not dually eligible</i>	84,228	119,209	\$3,844	\$1,029	110,517	128,741	\$3,789	\$1,195	-\$221	0.38	-\$636	\$195
IRF payments	<i>Dually eligible</i>	12,038	7,721	\$2,127	\$1,190	12,017	6,267	\$1,830	\$1,514	-\$620	0.10	-\$1,233	-\$7
	<i>Not dually eligible</i>	84,228	119,209	\$1,611	\$638	110,517	128,741	\$1,431	\$909	-\$450	0.09	-\$888	-\$12
Readmission payments	<i>Dually eligible</i>	12,038	7,721	\$1,249	\$1,256	12,017	6,267	\$1,230	\$1,240	-\$3	0.98	-\$260	\$254
	<i>Not dually eligible</i>	84,228	119,209	\$931	\$732	110,517	128,741	\$797	\$741	-\$143	0.15	-\$305	\$19
Post episode payments	<i>Dually eligible</i>	12,038	7,721	\$1,651	\$2,091	12,017	6,267	\$1,644	\$1,984	\$101	0.19	-\$26	\$228
	<i>Not dually eligible</i>	84,228	119,209	\$1,077	\$1,442	110,517	128,741	\$1,071	\$1,445	-\$9	0.71	-\$51	\$32

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

CJR = Comprehensive Care for Joint Replacement; DDD = difference-in-differences-in-differences; DiD = difference-in-differences; HH = home health; IRF = inpatient rehabilitation facility; LCI = lower confidence interval; N = number; PY = performance year; SNF = skilled nursing facility; UCI = upper confidence interval.

* The sample for this estimate failed tests for parallel trends in the baseline period.

Source: CJR evaluation team analysis of Medicare claims and enrollment data for episodes initiated between January 1, 2012, and December 31, 2014, that ended between April 1, 2012, and March 31, 2015 (baseline) and episodes that ended between October 1, 2021, and December 31, 2024 (PY 6–PY 8 intervention).

Exhibit D-13: Difference in impact for total episode payment outcomes, PY 6–PY 8, elective patients who were dually eligible

Measure	Population	PY 6–PY 8 impact (DiD)	Difference in impact (DDD)	p-value	90% LCI	90% UCI
Total episode payments	<i>Dually eligible</i>	-\$1,513	-\$370	0.34	-\$1,019	\$278
	<i>Not dually eligible</i>	-\$1,142				
HH payments	<i>Dually eligible</i>	\$291	\$125	0.41	-\$122	\$372
	<i>Not dually eligible</i>	\$166				
SNF payments	<i>Dually eligible</i>	-\$861	-\$641	0.10	-\$1,275	-\$6
	<i>Not dually eligible</i>	-\$221				
IRF payments	<i>Dually eligible</i>	-\$620	-\$170	0.43	-\$522	\$182
	<i>Not dually eligible</i>	-\$450				
Readmission payments	<i>Dually eligible</i>	-\$3	\$140	0.26	-\$65	\$345
	<i>Not dually eligible</i>	-\$143				
Post episode payments	<i>Dually eligible</i>	\$101	\$110	0.15	-\$16	\$237
	<i>Not dually eligible</i>	-\$9				

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

CJR = Comprehensive Care for Joint Replacement; DiD = difference-in-differences; DDD = difference-in-difference-in-differences; HH = home health; IRF = inpatient rehabilitation facility; LCI = lower confidence interval; PY = performance year; SNF = skilled nursing facility; UCI = upper confidence interval.

* The sample for this estimate failed tests for parallel trends in the baseline period.

Source: CJR evaluation team analysis of Medicare claims and enrollment data for episodes initiated between January 1, 2012, and December 31, 2014, that ended between April 1, 2012, and March 31, 2015 (baseline) and episodes that ended between October 1, 2021, and December 31, 2024 (PY 6–PY 8 intervention).

Exhibit D-14: Quality outcomes, PY 6–PY 8, for elective patients who were dually eligible

Measure	Population	CJR				Control				PY 6–PY 8 impact (DiD)	p-value	90% LCI	90% UCI
		Baseline episodes (N)	PY 6–PY 8 episodes (N)	Baseline mean	PY 6–PY 8 mean	Baseline episodes (N)	PY 6–PY 8 episodes (N)	Baseline mean	PY 6–PY 8 mean				
Mortality rate	<i>Dually eligible</i>	12,114	7,738	0.8%	0.7%	12,106	6,294	0.8%	0.8%	-0.07 pp	0.64	-0.32 pp	0.18 pp
	<i>Not dually eligible</i>	84,669	119,519	0.5%	0.4%	111,059	129,114	0.5%	0.4%	-0.04 pp	0.34	-0.12 pp	0.03 pp
ED use	<i>Dually eligible</i>	12,027	7,715	18.8%	17.7%	12,011	6,262	19.0%	16.6%	1.36 pp	0.33	-0.94 pp	3.65 pp
	<i>Not dually eligible</i>	84,217	119,198	11.3%	11.6%	110,505	128,730	10.6%	10.9%	0.10 pp	0.77	-0.48 pp	0.69 pp
Readmission rate	<i>Dually eligible</i>	12,027	7,715	11.3%	8.5%	12,011	6,262	10.8%	8.1%	-0.06 pp	0.93	-1.29 pp	1.17 pp
	<i>Not dually eligible</i>	84,217	119,198	7.8%	5.2%	110,505	128,730	7.4%	5.2%	-0.43 pp	0.16	-0.93 pp	0.07 pp

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

CJR = Comprehensive Care for Joint Replacement; DDD = difference-in-differences-in-differences; DiD = difference-in-differences; ED = emergency department; HH = home health; IRF = inpatient rehabilitation facility; LCI = lower confidence interval; N = number; PY = performance year; SNF = skilled nursing facility; UCI = upper confidence interval.

* The sample for this estimate failed tests for parallel trends in the baseline period.

Source: CJR evaluation team analysis of Medicare claims and enrollment data for episodes initiated between January 1, 2012, and December 31, 2014, that ended between April 1, 2012, and March 31, 2015 (baseline) and episodes that ended between October 1, 2021, and December 31, 2024 (PY 6–PY 8 intervention).

Exhibit D-15: Difference in quality outcomes, PY 6–PY 8, elective patients who were dually eligible

Measure	Impact on patients who were dually eligible (DiD; pp)	Impact on patients who were not dually eligible (DiD; pp)	Difference in impact (DDD; pp)	p-value	90% LCI (pp)	90% UCI (pp)
Mortality rate	-0.07	-0.04	-0.03	0.87	-0.29	0.24
ED use	1.36	0.10	1.25	0.33	-0.88	3.38
Readmission rate	-0.06	-0.43	0.36	0.61	-0.80	1.53

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

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

* The sample for this estimate failed tests for parallel trends in the baseline period.

Source: CJR evaluation team analysis of Medicare claims and enrollment data for episodes initiated between January 1, 2012, and December 31, 2014, that ended between April 1, 2012, and March 31, 2015 (baseline) and episodes that ended between October 1, 2021, and December 31, 2024 (PY 6–PY 8 intervention).

Exhibit D-16: First PAC outcomes, PY 6–PY 8, for elective patients who were dually eligible

Measure	Population	CJR				Control				PY 6–PY 8 impact (DiD)	p-value	90% LCI	90% UCI
		Baseline episodes (N)	PY 6–PY 8 episodes (N)	Baseline mean	PY 6–PY 8 mean	Baseline episodes (N)	PY 6–PY 8 episodes (N)	Baseline mean	PY 6–PY 8 mean				
First PAC IRF	<i>Dually eligible</i>	11,999	7,697	13.9%	4.7%	11,972	6,251	12.0%	7.0%	-4.12 pp	0.14	-8.72 pp	0.47 pp
	<i>Not dually eligible</i>	83,880	118,979	11.3%	2.3%	110,317	128,447	10.2%	4.3%	-3.08 pp	0.12	-6.37 pp	0.20 pp
First PAC SNF	<i>Dually eligible</i>	11,999	7,697	50.6%	19.0%	11,972	6,251	45.3%	15.9%	-2.27 pp	0.46	-7.32 pp	2.79 pp
	<i>Not dually eligible</i>	83,880	118,979	37.4%	6.8%	110,317	128,447	36.0%	7.3%	-1.79 pp	0.47	-5.84 pp	2.27 pp
First PAC HH	<i>Dually eligible</i>	11,999	7,697	30.2%	55.1%	11,972	6,251	33.0%	45.8%	12.08 pp	<0.01	4.98 pp	19.18 pp
	<i>Not dually eligible</i>	83,880	118,979	41.8%	61.1%	110,317	128,447	38.4%	44.9%	12.75 pp	0.03	3.29 pp	22.22 pp
First PAC OP PT/OT	<i>Dually eligible</i>	11,999	7,697	2.8%	7.3%	11,972	6,251	4.0%	10.7%	-2.18 pp	0.13	-4.56 pp	0.19 pp
	<i>Not dually eligible</i>	83,880	118,979	2.9%	7.3%	110,317	128,447	4.6%	11.7%	-2.66 pp	0.23	-6.34 pp	1.01 pp
First PAC home without HH or OP PT/OT	<i>Dually eligible</i>	11,999	7,697	2.5%	13.9%	11,972	6,251	5.7%	20.5%	-3.50 pp	0.21	-8.13 pp	1.13 pp
	<i>Not dually eligible</i>	83,880	118,979	6.7%	22.4%	110,317	128,447	10.8%	31.7%	-5.22 pp	0.12	-10.76 pp	0.33 pp

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

CJR = Comprehensive Care for Joint Replacement; DDD = difference-in-differences-in-differences; DiD = difference-in-differences; HH = home health; IRF = inpatient rehabilitation facility; LCI = lower confidence interval; N = number; OP = outpatient; OT = occupational therapy; PAC = post-acute care; pp = percentage point; PT = physical therapy; PY = performance year; SNF = skilled nursing facility; UCI = upper confidence interval.

* The sample for this estimate failed tests for parallel trends in the baseline period.

Source: CJR evaluation team analysis of Medicare claims and enrollment data for episodes initiated between January 1, 2012, and December 31, 2014, that ended between April 1, 2012, and March 31, 2015 (baseline) and episodes that ended between October 1, 2021, and December 31, 2024 (PY 6–PY 8 intervention).

Exhibit D-17: Difference in quality outcomes, PY 6–PY 8, elective patients who were dually eligible

Measure	Impact on patients who were dually eligible (DiD; pp)	Impact on patients who were not dually eligible (DiD; pp)	Difference in impact (DDD; pp)	p-value	90% LCI (pp)	90% UCI (pp)
First PAC IRF	-4.12	-3.08	-1.04	0.49	-3.53	1.45
First PAC SNF	-2.27	-1.79	-0.48	0.78	-3.25	2.29
First PAC HH	12.08	12.75	-0.68	0.83	-5.95	4.60
First PAC OP PT/OT	-2.18	-2.66	0.48	0.74	-1.88	2.84
First PAC home without HH or OP PT/OT	-3.50	-5.22	1.72	0.59	-3.46	6.89

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

CJR = Comprehensive Care for Joint Replacement; DDD = difference-in-differences-in-differences; DiD = difference-in-differences; HH = home health; IRF = inpatient rehabilitation facility; LCI = lower confidence interval; OP = outpatient; OT = occupational therapy; PAC = post-acute care; pp = percentage point; PT = physical therapy; PY = performance year; SNF = skilled nursing facility; UCI = upper confidence interval.

* The sample for this estimate failed tests for parallel trends in the baseline period.

Source: CJR evaluation team analysis of Medicare claims and enrollment data for episodes initiated between January 1, 2012, and December 31, 2014, that ended between April 1, 2012, and March 31, 2015 (baseline) and episodes that ended between October 1, 2021, and December 31, 2024 (PY 6–PY 8 intervention).

Exhibit D-18: Utilization outcomes, PY 6–PY 8, elective patients who were dually eligible

Measure	Population	CJR				Control				PY 6–PY 8 impact (DiD)	p-value	90% LCI	90% UCI
		Baseline episodes (N)	PY 6–PY 8 episodes (N)	Baseline mean	PY 6–PY 8 mean	Baseline episodes (N)	PY 6–PY 8 episodes (N)	Baseline mean	PY 6–PY 8 mean				
SNF days	<i>Dually eligible</i>	6,216	1,749	27.1	27.1	5,238	1,081	25.51	26.53	-1.03	0.44	-3.25	1.19
	<i>Not dually eligible</i>	31,956	9,442	19.2	17.7	39,796	9,090	19.8	18.6	-0.29	0.70	-1.54	0.96
IRF days	<i>Dually eligible</i>	1,730	447	10.6	11.5	1,538	468	10.6	11.7	-0.16	0.68	-0.80	0.48
	<i>Not dually eligible</i>	9,595	3,691	10.2	11.4	10,818	5,685	10.0	11.1	0.11	0.69	-0.34	0.55

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

CJR = Comprehensive Care for Joint Replacement; DDD = difference-in-differences-in-differences; DiD = difference-in-differences; HH = home health; IRF = inpatient rehabilitation facility; LCI = lower confidence interval; N = number; PY = performance year; SNF = skilled nursing facility; UCI = upper confidence interval.

* The sample for this estimate failed tests for parallel trends in the baseline period.

Source: CJR evaluation team analysis of Medicare claims and enrollment data for episodes initiated between January 1, 2012, and December 31, 2014, that ended between April 1, 2012, and March 31, 2015 (baseline) and episodes that ended between October 1, 2021, and December 31, 2024 (PY 6–PY 8 intervention).

Exhibit D-19: Difference in impact utilization outcomes, PY 6–PY 8, elective patients who were dually eligible

Measure	Impact on patients who were dually eligible (DiD)	Impact on patients who were not dually eligible (DiD)	Difference in impact (DDD)	p-value	90% LCI	90% UCI
SNF days	-1.03	-0.29	-0.74	0.60	-3.07	1.58
IRF days	-0.16	0.11	-0.27	0.47	-0.88	0.35

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

CJR = Comprehensive Care for Joint Replacement; DDD = difference-in-differences-in-differences; DiD = difference-in-differences; HH = home health; IRF = inpatient rehabilitation facility; LCI = lower confidence interval; PY = performance year; SNF = skilled nursing facility; UCI = upper confidence interval.

* The sample for this estimate failed tests for parallel trends in the baseline period.

Source: CJR evaluation team analysis of Medicare claims and enrollment data for episodes initiated between January 1, 2012, and December 31, 2014, that ended between April 1, 2012, and March 31, 2015 (baseline) and episodes that ended between October 1, 2021, and December 31, 2024 (PY 6–PY 8 intervention).

Exhibit D-20: Episode payment outcomes, PY 6–PY 8, patients who were dually eligible with a fracture

Measure	Population	CJR				Control				PY 6–PY 8 impact (DiD)	p-value	90% LCI	90% UCI
		Baseline episodes (N)	PY 6–PY 8 episodes (N)	Baseline mean	PY 6–PY 8 mean	Baseline episodes (N)	PY 6–PY 8 episodes (N)	Baseline mean	PY 6–PY 8 mean				
Total episode payments	<i>Dually eligible</i>	3,622	2,637	\$51,096	\$58,302	3,056	1,785	\$51,203	\$56,669	\$1,740	0.21	-\$564	\$4,043
	<i>Not dually eligible</i>	14,637	13,713	\$46,407	\$48,997	15,883	14,319	\$46,246	\$49,395	-\$560	0.48	-\$1,887	\$768
HH payments	<i>Dually eligible</i>	3,622	2,637	\$1,698	\$1,792	3,056	1,785	\$1,744	\$1,805	\$33	0.76	-\$142	\$209
	<i>Not dually eligible</i>	14,637	13,713	\$2,624	\$2,655	15,883	14,319	\$2,493	\$2,559	-\$35	0.64	-\$158	\$88
SNF payments	<i>Dually eligible</i>	3,622	2,637	\$21,495	\$24,315	3,056	1,785	\$22,640	\$23,723	\$1,738	0.14	-\$205	\$3,680
	<i>Not dually eligible</i>	14,637	13,713	\$15,554	\$13,981	15,883	14,319	\$16,422	\$13,819	\$1,029	0.15	-\$153	\$2,212
IRF payments	<i>Dually eligible</i>	3,622	2,637	\$3,874	\$4,683	3,056	1,785	\$2,993	\$4,534	-\$732	0.38	-\$2,111	\$647
	<i>Not dually eligible</i>	14,637	13,713	\$5,748	\$7,058	15,883	14,319	\$5,263	\$7,846	-\$1,272	0.12	-\$2,607	\$62
Readmission payments	<i>Dually eligible</i>	3,622	2,637	\$3,129	\$3,623	3,056	1,785	\$3,066	\$3,284	\$276	0.55	-\$489	\$1,042
	<i>Not dually eligible</i>	14,637	13,713	\$2,499	\$2,612	15,883	14,319	\$2,334	\$2,612	-\$165	0.23	-\$394	\$63
Post episode payments	<i>Dually eligible</i>	3,622	2,637	\$4,097	\$5,324	3,056	1,785	\$4,399	\$4,620	\$1,005	<0.01	\$463	\$1,548
	<i>Not dually eligible</i>	14,637	13,713	\$3,238	\$3,667	15,883	14,319	\$3,377	\$3,531	\$275	0.05	\$41	\$508

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

CJR = Comprehensive Care for Joint Replacement; DDD = difference-in-difference-in-differences; DiD = difference-in-differences; HH = home health; IRF = inpatient rehabilitation facility; LCI = lower confidence interval; N = number; PY = performance year; SNF = skilled nursing facility; UCI = upper confidence interval.

* The sample for this estimate failed tests for parallel trends in the baseline period.

Source: CJR evaluation team analysis of Medicare claims and enrollment data for episodes initiated between January 1, 2012, and December 31, 2014, that ended between April 1, 2012, and March 31, 2015 (baseline) and episodes that ended between October 1, 2021, and December 31, 2024 (PY 6–PY 8 intervention).

Exhibit D-21: Difference in impact for total episode payment outcomes, PY 6–PY 8, patients who were dually eligible with a fracture

Measure	Population	PY 6–PY 8 impact (DiD)	Difference in impact (DDD)	p-value	90% LCI	90% UCI
Total episode payments	<i>Dually eligible</i>	\$1,740	\$2,299	0.06	\$257	\$4,342
	<i>Not dually eligible</i>	-\$560				
HH payments	<i>Dually eligible</i>	\$33	\$68	0.48	-\$92	\$228
	<i>Not dually eligible</i>	-\$35				
SNF payments	<i>Dually eligible</i>	\$1,738	\$708	0.41	-\$709	\$2,126
	<i>Not dually eligible</i>	\$1,029				
IRF payments	<i>Dually eligible</i>	-\$732	\$540	0.38	-\$463	\$1,543
	<i>Not dually eligible</i>	-\$1,272				
Readmission payments	<i>Dually eligible</i>	\$276	\$442	0.33	-\$301	\$1,184
	<i>Not dually eligible</i>	-\$165				
Post episode payments	<i>Dually eligible</i>	\$1,005	\$730	0.02	\$197	\$1,264
	<i>Not dually eligible</i>	\$275				

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

CJR = Comprehensive Care for Joint Replacement; DDD = difference-in-difference-in-differences; DiD = difference-in-differences; HH = home health; IRF = inpatient rehabilitation facility; LCI = lower confidence interval; PY = performance year; SNF = skilled nursing facility; UCI = upper confidence interval.

* The sample for this estimate failed tests for parallel trends in the baseline period.

Source: CJR evaluation team analysis of Medicare claims and enrollment data for episodes initiated between January 1, 2012, and December 31, 2014, that ended between April 1, 2012, and March 31, 2015 (baseline) and episodes that ended between October 1, 2021, and December 31, 2024 (PY 6–PY 8 intervention).

Exhibit D-22: Quality outcomes, PY 6–PY 8, for patients who were dually eligible with a fracture

Measure	Population	CJR				Control				PY 6–PY 8 impact (DiD)	p-value	90% LCI	90% UCI
		Baseline episodes (N)	PY 6–PY 8 episodes (N)	Baseline mean	PY 6–PY 8 mean	Baseline episodes (N)	PY 6–PY 8 episodes (N)	Baseline mean	PY 6–PY 8 mean				
Mortality rate	<i>Dually eligible</i>	4,077	2,938	13.1%	12.5%	3,538	1,999	15.7%	14.3%	0.82 pp	0.52	-1.28 pp	2.92 pp
	<i>Not dually eligible</i>	16,555	15,276	12.7%	11.6%	17,977	15,956	12.7%	12.1%	-0.49 pp	0.33	-1.30 pp	0.33 pp
ED use	<i>Dually eligible</i>	3,620	2,632	18.4%	20.4%	3,055	1,781	20.3%	21.6%	0.63 pp	0.76	-2.79 pp	4.05 pp
	<i>Not dually eligible</i>	14,636	13,704	17.8%	18.3%	15,879	14,317	16.6%	18.4%	-1.31 pp	0.10	-2.63 pp	0.01 pp
Readmission rate	<i>Dually eligible</i>	3,620	2,632	23.8%	20.5%	3,055	1,781	23.6%	17.9%	2.39 pp	0.23	-0.89 pp	5.68 pp
	<i>Not dually eligible</i>	14,636	13,704	20.7%	15.6%	15,879	14,317	19.7%	16.3%	-1.68 pp	0.08	-3.25 pp	-0.11 pp

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

CJR = Comprehensive Care for Joint Replacement; DDD = difference-in-difference-in-differences; DiD = difference-in-differences; LCI = lower confidence interval; N = number; PY = performance year; UCI = upper confidence interval.

* The sample for this estimate failed tests for parallel trends in the baseline period.

Source: CJR evaluation team analysis of Medicare claims and enrollment data for episodes initiated between January 1, 2012, and December 31, 2014, that ended between April 1, 2012, and March 31, 2015 (baseline) and episodes that ended between October 1, 2021, and December 31, 2024 (PY 6–PY 8 intervention).

Exhibit D-23: Difference in quality outcomes, PY 6–PY 8, patients who were dually eligible with a fracture

Measure	Impact on patients who were dually eligible (DiD)	Impact on patients who were not dually eligible (DiD)	Difference in impact (DDD)	p-value	90% LCI	90% UCI
Mortality rate	0.82 pp	-0.49 pp	1.31 pp	0.26	-0.58 pp	3.19 pp
ED use	0.63 pp	-1.31 pp	1.94 pp	0.25	-0.84 pp	4.71 pp
Readmission rate	2.39 pp	-1.68 pp	4.07 pp	0.01	1.45 pp	6.70 pp

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

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

* The sample for this estimate failed tests for parallel trends in the baseline period.

Source: CJR evaluation team analysis of Medicare claims and enrollment data for episodes initiated between January 1, 2012, and December 31, 2014, that ended between April 1, 2012, and March 31, 2015 (baseline) and episodes that ended between October 1, 2021, and December 31, 2024 (PY 6–PY 8 intervention).

Exhibit D-24: First PAC outcomes, PY 6–PY 8, for patients who were dually eligible with a fracture

Measure	Population	CJR				Control				PY 6–PY 8 impact (DiD)	p-value	90% LCI	90% UCI
		Baseline episodes (N)	PY 6–PY 8 episodes (N)	Baseline mean	PY 6–PY 8 mean	Baseline episodes (N)	PY 6–PY 8 episodes (N)	Baseline mean	PY 6–PY 8 mean				
First PAC IRF	<i>Dually eligible</i>	3,547	2,606	19.2%	16.6%	3,004	1,759	15.2%	16.8%	-4.26 pp	0.18	-9.47 pp	0.95 pp
	<i>Not dually eligible</i>	14,525	13,633	29.4%	26.4%	15,694	14,207	28.0%	30.9%	-5.83 pp	0.13	-12.23 pp	0.57 pp
First PAC SNF	<i>Dually eligible</i>	3,547	2,606	72.5%	68.5%	3,004	1,759	76.0%	72.1%	-0.03 pp	0.99	-5.12 pp	5.06 pp
	<i>Not dually eligible</i>	14,525	13,633	63.1%	54.6%	15,694	14,207	63.9%	53.1%	2.20 pp	0.43	-2.41 pp	6.81 pp
First PAC HH	<i>Dually eligible</i>	3,547	2,606	4.7%	10.3%	3,004	1,759	5.7%	6.9%	4.50 pp	<0.01	2.34 pp	6.66 pp
	<i>Not dually eligible</i>	14,525	13,633	5.5%	15.2%	15,694	14,207	6.0%	12.3%	3.36 pp	0.01	1.11 pp	5.61 pp
First PAC OP PT/OT	<i>Dually eligible</i>	3,547	2,606	1.3%	2.5%	3,004	1,759	1.7%	2.3%	0.59 pp	0.19	-0.16 pp	1.34 pp
	<i>Not dually eligible</i>	14,525	13,633	1.4%	2.4%	15,694	14,207	1.4%	2.2%	0.13 pp	0.77	-0.62 pp	0.89 pp
First PAC home without HH or OP PT/OT	<i>Dually eligible</i>	3,547	2,606	2.3%	2.1%	3,004	1,759	1.4%	1.9%	-0.80 pp	0.13	-1.67 pp	0.08 pp
	<i>Not dually eligible</i>	14,525	13,633	0.5%	1.5%	15,694	14,207	0.7%	1.5%	0.14 pp	0.60	-0.29 pp	0.56 pp

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

CJR = Comprehensive Care for Joint Replacement; HH = home health; IRF = inpatient rehabilitation facility; DDD = difference-in-difference-in-differences; DiD = difference-in-differences; LCI = lower confidence interval; N = number; OP = outpatient; OT = occupational therapy; PAC = post-acute care; pp = percentage point; PT = physical therapy; PY = performance year; SNF = skilled nursing facility; UCI = upper confidence interval.

* The sample for this estimate failed tests for parallel trends in the baseline period.

Source: CJR evaluation team analysis of Medicare claims and enrollment data for episodes initiated between January 1, 2012, and December 31, 2014, that ended between April 1, 2012, and March 31, 2015 (baseline) and episodes that ended between October 1, 2021, and December 31, 2024 (PY 6–PY 8 intervention).

Exhibit D-25: Difference in first PAC outcomes, PY 6–PY 8, patients who were dually eligible with a fracture

Measure	Impact on patients who were dually eligible (DiD)	Impact on patients who were not dually eligible (DiD)	Difference in impact (DDD)	p-value	90% LCI	90% UCI
First PAC IRF	-4.26 pp	-5.83 pp	1.57 pp	0.51	-2.39 pp	5.53 pp
First PAC SNF	-0.03 pp	2.20 pp	-2.23 pp	0.33	-6.02 pp	1.56 pp
First PAC HH	4.50 pp	3.36 pp	1.14 pp	0.39	-1.03 pp	3.31 pp
First PAC OP PT/OT	0.59 pp	0.13 pp	0.46 pp	0.48	-0.60 pp	1.51 pp
First PAC home without HH or OP PT/OT	-0.80 pp	0.14 pp	-0.93 pp	0.14	-1.97 pp	0.11 pp

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

CJR = Comprehensive Care for Joint Replacement; HH = home health; IRF = inpatient rehabilitation facility; DDD = difference-in-difference-in-differences; DiD = difference-in-differences; LCI = lower confidence interval; OP = outpatient; OT = occupational therapy; PAC = post-acute care; pp = percentage point; PT = physical therapy; PY = performance year; SNF = skilled nursing facility; UCI = upper confidence interval.

*The sample for this estimate failed tests for parallel trends in the baseline period.

Source: CJR evaluation team analysis of Medicare claims and enrollment data for episodes initiated between January 1, 2012, and December 31, 2014, that ended between April 1, 2012, and March 31, 2015 (baseline) and episodes that ended between October 1, 2021, and December 31, 2024 (PY 6–PY 8 intervention).

Exhibit D-26: Utilization outcomes, PY 6–PY 8, patients who were dually eligible with a fracture

Measure	Population	CJR				Control				PY 6–PY 8 impact (DiD)	p-value	90% LCI	90% UCI
		Baseline episodes (N)	PY 6–PY 8 episodes (N)	Baseline mean	PY 6–PY 8 mean	Baseline episodes (N)	PY 6–PY 8 episodes (N)	Baseline mean	PY 6–PY 8 mean				
SNF days	<i>Dually eligible</i>	2,867	1,980	51.1	48.4	2,424	1,293	52.4	47.1	2.47	0.11	-0.07	5.01
	<i>Not dually eligible</i>	10,503	8,092	41.0	34.7	11,050	7,819	43.0	35.5	1.26	0.29	-0.70	3.21
IRF days	<i>Dually eligible</i>	741	441	14.6	14.2	527	332	14.3	14.1	-0.15	0.80	-1.17	0.87
	<i>Not dually eligible</i>	4,349	3,860	14.1	13.7	4,961	4,933	13.6	13.4	-0.16	0.65	-0.75	0.43

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

CJR = Comprehensive Care for Joint Replacement; DDD = difference-in-difference-in-differences; DiD = difference-in-differences; HH = home health; IRF = inpatient rehabilitation facility; LCI = lower confidence interval; N = number; PY = performance year; SNF= skilled nursing facility; UCI = upper confidence interval.

* The sample for this estimate failed tests for parallel trends in the baseline period.

Source: CJR evaluation team analysis of Medicare claims and enrollment data for episodes initiated between January 1, 2012, and December 31, 2014, that ended between April 1, 2012, and March 31, 2015 (baseline) and episodes that ended between October 1, 2021, and December 31, 2024 (PY 6–PY 8 intervention).

Exhibit D-27: Difference in impact utilization outcomes, PY 6–PY 8, patients who were dually eligible

Measure	Impact on patients who were dually eligible (DiD)	Impact on patients who were not dually eligible (DiD)	Difference in impact (DDD)	p-value	90% LCI	90% UCI
SNF days	2.47	1.26	1.22	0.29	-0.69	3.13
IRF days	-0.15	-0.16	0.01	0.99	-1.04	1.06

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

CJR = Comprehensive Care for Joint Replacement; DiD = difference-in-differences; DDD = difference-in-difference-in-differences; HH = home health; IRF = inpatient rehabilitation facility; LCI = lower confidence interval; PY = performance year; SNF= skilled nursing facility; UCI = upper confidence interval.

* The sample for this estimate failed tests for parallel trends in the baseline period.

Source: CJR evaluation team analysis of Medicare claims and enrollment data for episodes initiated between January 1, 2012, and December 31, 2014, that ended between April 1, 2012, and March 31, 2015 (baseline) and episodes that ended between October 1, 2021, and December 31, 2024 (PY 6–PY 8 intervention).

B. Medicare Program Savings

The results presented in this section correspond to the findings presented in **Chapter IV: Medicare Program Savings**.

Exhibit D-28: Medicare program savings, PY 6, all LEJR

Time period	Value	90% LCI	90% UCI
Reduction in non-standardized paid amounts per episode	\$920	-\$15	\$1,856
Reconciliation payments per episode	-\$78	N/A	N/A
Medicare savings per episode*	\$998	\$63	\$1,934
Number of PY 6 episodes	53,326	N/A	N/A
Aggregate Medicare savings*	\$53,240,346	\$3,353,216	\$103,127,471

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

CJR = Comprehensive Care for Joint Replacement; LCI = lower confidence interval; LEJR = lower extremity joint replacement; N/A = not applicable; PY = performance year; UCI = upper confidence interval.

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

Exhibit D-29: Medicare program savings, PY 7, all LEJR

Time period	Value	90% LCI	90% UCI
Reduction in non-standardized paid amounts per episode	\$878	-\$170	\$1,927
Reconciliation payments per episode	-\$404	N/A	N/A
Medicare savings per episode*	\$1,282	\$234	\$2,331
Number of PY 7 episodes	45,415	N/A	N/A
Aggregate Medicare savings*	\$58,239,841	\$10,613,656	\$105,866,041

Notes: Reductions in nonstandardized paid amounts are based on estimates from a DiD model of per-episode standardized paid amounts that have been multiplied by –1 and converted to nonstandardized amounts. We do not report confidence intervals for reconciliation payments per episode and number of PY 7 episodes because these were not estimated but observed with certainty. We calculated aggregate Medicare savings by multiplying Medicare savings per episode by the number of PY 7 episodes, and Medicare savings per episode equals the estimated reduction in nonstandardized paid amounts per episode minus the average reconciliation payments per episode. Because reconciliation payments per episode depend on realized payment reductions, the asterisked intervals are not truly confidence intervals but ranges based on the confidence interval.

CJR = Comprehensive Care for Joint Replacement; LCI = lower confidence interval; LEJR = lower extremity joint replacement; N/A = not applicable; PY = performance year; UCI = upper confidence interval.

Source: CJR evaluation team analysis of Medicare claims and enrollment data for episodes initiated in 2012 through 2014 that ended between April 2012 and March 2015 (baseline) and episodes that ended between January 1, 2023, and December 31, 2023 (PY 7 intervention) as well as CJR payment contractor data for CJR participant hospitals in PY 7.

Exhibit D-30: Medicare program savings, PY 8, all LEJR

Time period	Value	90% LCI	90% UCI
Reduction in non-standardized paid amounts per episode	\$1,059	-\$130	\$2,248
Reconciliation payments per episode	-\$471	N/A	N/A
Medicare savings per episode*	\$1,530	\$341	\$2,719
Number of PY 8 episodes	44,628	N/A	N/A
Aggregate Medicare savings*	\$68,272,608	\$15,201,027	\$121,344,180

Notes: Reductions in nonstandardized paid amounts are based on estimates from a DiD model of per-episode standardized paid amounts that have been multiplied by –1 and converted to nonstandardized amounts. We do not report confidence intervals for reconciliation payments per episode and number of PY 8 episodes because these were not estimated but observed with certainty. We calculated aggregate Medicare savings by multiplying Medicare savings per episode by the number of PY 8 episodes, and Medicare savings per episode equals the estimated reduction in nonstandardized paid amounts per episode minus the average reconciliation payments per episode. Because reconciliation payments per episode depend on realized payment reductions, the asterisked intervals are not truly confidence intervals but ranges based on the confidence interval.

CJR = Comprehensive Care for Joint Replacement; LCI = lower confidence interval; LEJR = lower extremity joint replacement; M = million; N/A = not applicable; PY = performance year; UCI = upper confidence interval.

Source: CJR evaluation team analysis of Medicare claims and enrollment data for episodes initiated in 2012 through 2014 that ended between April 2012 and March 2015 (baseline) and episodes that ended between January 1, 2024, and December 31, 2024 (PY 8 intervention) as well as CJR payment contractor data for CJR participant hospitals in PY 8.