LEWINGROUP®

CMS Comprehensive Care for Joint Replacement Model: Performance Year 2 Evaluation Report

Second Annual Report

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

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

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

June 2019

CMS Comprehensive Care for Joint Replacement Model: Performance Year 2 Evaluation Report

Second Annual Report

Prepared for:

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

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

Submitted by:

The Lewin Group, Inc.

June 2019

Table of Contents

EXECUTIVE SUMMARY **1**

I. INTRODUCTION **11**

A. The CJR Model 11

B. Evaluation Conceptual Framework 13

II. RESULTS **17**

A. What was the impact of the CJR model on average episode payments? 17

1. Key findings *17*

2. Methods *17*

3. Results *18*

B. How much did the Medicare program save (or lose) due to the CJR model after accounting for reconciliation payments? 19

1. Key findings *19*

2. Methods *19*

3. Results *20*

C. What types of CJR participant hospitals did and did not receive reconciliation payments? 22

1. Key findings *23*

2. Methods *23*

3. Results *23*

D. What was the impact of the CJR model on service-level payments and service use during the episode? 28

1. Key finding *28*

2. Methods *28*

3. Results *29*

E. What was the impact of the CJR model on quality of care? 35

1. Key finding *36*

2. Methods *36*

3. Results *36*

F. What was the impact of the CJR model on functional status, pain, and care experiences? 37

1. Key findings *37*

2. Methods *37*

3. Results *38*

G. Did the CJR model result in any unintended consequences? 49
1. What was the impact of the CJR model on total market volume of elective LEJR discharges? *50*
2. Are there any indications that the CJR patient population was healthier in the intervention period than in the baseline period? *52*
3. What was the impact of the CJR model on payments in the 30 days following the episode? *57*
H. How did the impact of the CJR model vary for particular subgroups of MSAs, hospitals, and beneficiaries? 57
1. Key findings *59*
2. Methods *59*
3. Results *60*
I. What key factors and model features influenced hospitals' choice of response to the CJR model? 74
1. Key findings *74*
2. Methods *75*
3. Results *75*
J. What did CJR participant hospitals do to redesign care for their LEJR patients? 78
1. Key findings *78*
2. Methods *78*
3. Results *78*
K. How were relationships with orthopedic surgeons and PAC providers impacted by the CJR model? 86
1. Key findings *87*
2. Methods *87*
3. Results *87*

III. DISCUSSION AND CONCLUSION 92

A. Discussion 92
B. Considerations 93
C. Conclusion 94

List of Appendices

Appendix A: List of Acronyms & Glossary Terms A-1

Appendix B: CJR Programmatic Flexibilities, Including Financial Arrangements, Beneficiary Incentives, and Program Rule Waivers B-1

Appendix C: CJR Participant Hospitals Compared with All Other IPPS Hospitals C-1

Appendix D: Definitions of Hospital and Patient Characteristics D-1

Appendix E: Methodology E-1

Appendix F: Characteristics of CJR Participant Hospitals that did and did not Receive Reconciliation Payments F-1

Appendix G: Outcome Definitions G-1

Appendix H: Payment, Utilization, Quality, and Activities of Daily Living Results H-1

Appendix I: Patient Survey Questions I-1

Appendix J: Activities of Daily Living Sensitivity Analyses J-1

Appendix K: Patient Survey Results K-1

Appendix L: Change in Patient Characteristics L-1

Appendix M: Hospital Survey Questions M-1

Executive Summary

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

The second annual CJR model evaluation report presents findings from the first two performance years, which include episodes initiated on or after April 1, 2016 that ended by December 31, 2017. During this period, the model was mandatory for nearly all hospitals in 67 geographic areas, defined by Metropolitan Statistical Areas (MSAs).[3]

In the first two performance years, the CJR model achieved a statistically significant reduction in average episode payments due to reductions in institutional post-acute care (PAC) use. After accounting for reconciliation payments paid to participant hospitals, Medicare savings from the CJR model was estimated to be $17.4 million, although estimated savings ranged from a net loss of $41.2 million to savings of up to $75.9 million because of uncertainty around the estimated reduction in episode payments. Quality of care, as measured by the readmission rate, emergency department visits, and mortality, was maintained under the CJR model and patient surveys indicated that CJR and control patients had similar changes in functional status from before their surgery to after the episode. CJR participant hospital representatives we interviewed reported that they responded to the model by beginning discharge planning earlier, educating patients about discharge to less intensive PAC settings, and coordinating with PAC providers.

A. Introduction

CJR participant hospitals are financially accountable for the cost and quality of health care services for LEJR episodes of care. The CJR model is intended to reward participant hospitals for reducing episode payments and improving quality by coordinating care with the physicians, PAC providers, and other providers and clinicians involved in the episode. Through an annual reconciliation process, participant hospitals may earn reconciliation payments if they achieve cost and quality metrics or face repayments to Medicare if they do not. The CJR model has a mandatory,

---

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

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

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

randomized design in which hospitals in randomly selected MSAs were required to participate. Because of this design, a spectrum of hospitals with varying levels of infrastructure, care redesign experience, episode costs and utilization, and market positions are participating, which allows a broad test of the CJR model.

Episode definition. Under the CJR model, an LEJR episode of care begins with the hospitalization for the surgery and extends through the 90 days after hospital discharge. All Medicare-covered items and services provided during this period, with some exclusions, are included in the episode bundle.[4] All providers and suppliers involved in the episode continue to be paid under Medicare's fee-for-service (FFS) payment systems.

Annual reconciliation. After the end of each model performance year, CMS reconciles each participant hospital's LEJR episode payments against the hospital's quality-adjusted target price. The quality-adjusted target price is based on a discounted blend of the hospital's average historical episode payments and the region's average historical episode payments. During the first two performance years, two-thirds of the quality-adjusted target price is the hospital's average and one-third is the regional average. The discount to the quality-adjusted target price is intended to be Medicare's portion of the decrease in spending under the model. At reconciliation, the discount is adjusted based on the participant hospital's composite quality score. A lower discount is applied to the target price for participant hospitals with a higher quality score (resulting in a higher quality-adjusted target price) to encourage participant hospitals to focus on improving quality.

Hospitals with LEJR episode payments *below* their quality-adjusted target price and an "acceptable" or higher composite quality score receive a reconciliation payment. The reconciliation payment equals the difference between the quality-adjusted target price and actual episode payments, up to a stop-gain limit. Starting in performance year (PY) 2, hospitals with episode payments *above* their quality-adjusted target price repay Medicare the difference, subject to a stop-loss limit. In PY1, this repayment responsibility was forgiven to allow hospitals time to gain experience under the CJR model before implementation of two-sided risk.

Mandatory, randomized design. The mandatory, randomized design of the CJR model results in a diverse group of CJR participant hospitals that includes hospitals that might not voluntarily participate in an episode-based payment model. For the two performance years that are covered in this annual report, all acute care hospitals paid under the Medicare inpatient prospective payment system (IPPS), with few exceptions, in 67 randomly selected MSAs were required to participate. The mandatory MSAs were identified from 171 MSAs that were eligible for participation when the model design was finalized. MSAs were selected for participation using

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

eight sampling strata based on a median split of MSA population size and quartiles of average MSA historical episode payments.[5] An MSA's probability of selection increased with the payment quartiles in order to oversample high payment MSAs for participation in CJR. This was because of the belief that there is greater need and more opportunities for payment reductions in higher payment areas. Eligible MSAs that were not selected are a natural control group for evaluating the impact of the CJR model.

For more information about the CJR model, visit: https://innovation.cms.gov/initiatives/cjr

B. Results

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

During the first two performance years, the CJR model resulted in decreases in average standardized allowed amounts (payments) for LEJR episodes. Payments decreased by $997 more for CJR episodes than for control group episodes, or 3.7% from CJR baseline payments ($p<0.01$).[6] Average episode payments decreased for all subgroups of MSAs, hospitals, and episodes that we examined. Decreases in payments were due to shifts from more to less intensive PAC. CJR participant hospitals discharged a relatively smaller proportion of patients to an inpatient rehabilitation facility (IRF) and a relatively larger proportion of patients to a home health agency (HHA) than control group hospitals. Furthermore, CJR patients with a skilled nursing facility (SNF) stay spent relatively fewer days in a SNF than control group patients. These shifts in utilization resulted in statistically significant decreases in IRF and SNF payments, which drove the decrease in average episode payments.

Even though the CJR model resulted in a decrease in average episode payments, after accounting for the reconciliation payments to and repayments from participant hospitals, we estimate that

[5] Originally, 196 MSAs were identified as eligible for participation in the CJR model and the mandatory MSAs were randomly selected from this pool. CMS later identified 25 MSAs that were ineligible for selection after accounting for BPCI PGP participation.

[6] These results are based on the difference-in-differences statistical technique, which quantifies the impact of the CJR model by comparing changes in the outcome for CJR participant hospitals to changes for a control group from a baseline to the intervention period. To account for any differences between the CJR and control groups, we risk adjusted outcomes for hospital and patient characteristics, as well as geographic location. Payment outcomes are based on standardized Medicare allowed amounts. Standardizing payments removes wage adjustments and other Medicare payment adjustments and allowed amounts include beneficiary cost sharing.

Medicare savings due to the CJR model was $17.4 million. Because of uncertainty around the per episode payment decrease, however, the Medicare savings estimate ranges from Medicare losses of up to $41.2 million to savings of up to $75.9 million during the first two performance years of the CJR model.

Quality of care, as measured by the unplanned readmission rate, emergency department visits, and mortality, was maintained under the CJR model. Further, by the end of the episode, CJR and comparison patient survey respondents reported similar functional status gains and pain levels from before their hospitalization to after the end of the episode. For patients discharged to PAC, the proportion of CJR patients who improved their functional status during their PAC stay decreased relative to control patients. Orthopedic surgeons and other clinicians we interviewed and consulted were consistent in their view that home was the best place for most patients to recover.

The evaluation also examined whether CJR participant hospitals reduced episode payments through means other than redesigning care, or unintended consequences, which could result in higher Medicare program spending. For instance, CJR participant hospitals could provide LEJR to patients who would not have received LEJR in the absence of the model, delay services until after the end of the episode, or favor less complex patients who may be the least costly. The CJR model is designed to protect against these responses by including all hospitals in the MSA, the long episode definition, and other means, but they are still possible. We found no evidence that the CJR model was associated with an increase in the total market volume of LEJR episodes or that services were delayed until after the end of the episode. There is some evidence that the population of CJR patients receiving elective LEJR without major complications or comorbidities became healthier on average relative to the control group, which could indicate patient selection or induced demand. This preliminary finding will be the subject of further investigation.

CJR participant hospital representatives we interviewed described several factors that motivated the changes they made in response to the CJR model, including the opportunity to prepare for future bundled payment models. While hospitals reported considering the possibility of financial gains and losses under the CJR model, many stated that financial pressure was not the primary driver in their decisions about responding. In fact, they reported that the influence of CJR was often not distinguishable from other market factors that affected decisions about their orthopedic service lines. For hospitals that chose to make changes in response to the CJR model, nearly all implemented changes that were intended to alter PAC use and improve quality. In some instances, interviewees said that care redesign efforts were underway prior to the CJR model or were implemented because of factors unrelated to CJR. Interviewees reported a variety of approaches to care redesign that affected care prior to the surgery, during the inpatient hospital stay, and during the 90 days following discharge from the hospital. Many also indicated that they started or bolstered education and coordination efforts that targeted patients, hospital staff, orthopedic surgeons, and PAC providers.

Additional details about key findings are summarized under the main research questions addressed in this report.

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

- **Average episode payments decreased $997 more for CJR episodes than for control group episodes during the first two performance years of the CJR model ($p<0.01$).**[7] This relative reduction in payments equates to a 3.7% decrease in average episode payments for CJR episodes from the baseline.[8]

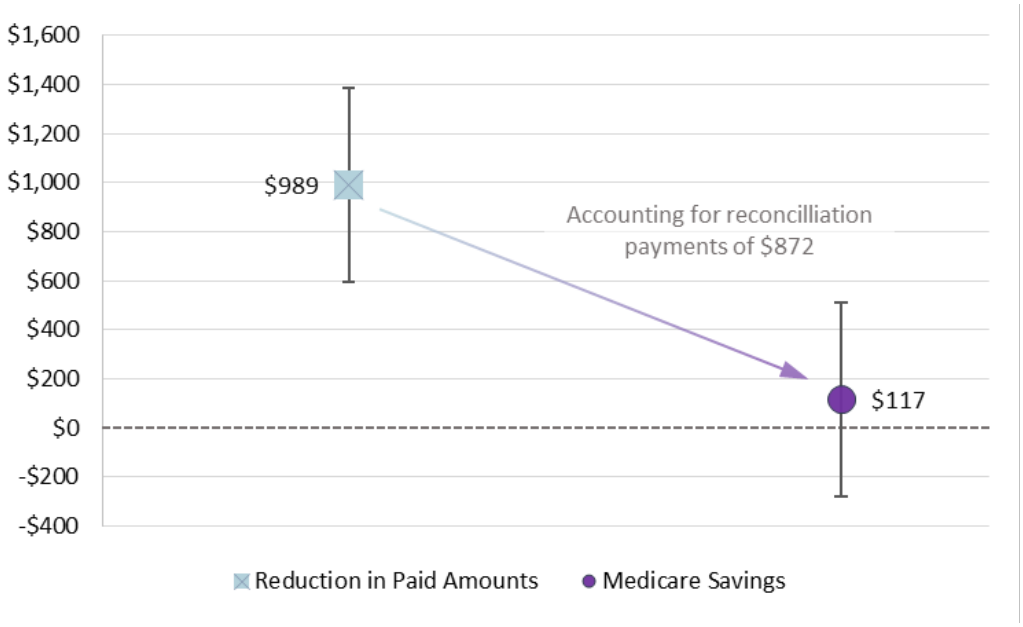
2. How much did the Medicare program save (or lose) due to the CJR model after accounting for reconciliation payments?

- **The CJR model likely resulted in savings to Medicare, but there is a wide range around the estimated savings.** We used non-standardized allowed amounts, that is, the payment from Medicare to providers that includes geographic and other payment adjustments and excludes beneficiary cost sharing, to estimate the savings to the Medicare program. Total non-standardized allowed amounts decreased by $146.3 million, with a 90% confidence interval that ranged from a decrease of $87.8 million to a decrease of $204.8 million. After accounting for $128.9 million in reconciliation payments made to and repayments from CJR participant hospitals, estimated savings to Medicare was $17.4 million ($117 per episode), ranging from losses of up to $41.2 million ($278 per episode) to savings of up to $75.9 million ($513 per episode) (Exhibit 1). While the CJR model reduced average episode payments, due to the wide range around the estimated decrease we cannot conclude with statistical certainty that the CJR model resulted in savings to Medicare in its first two performance years.

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

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

Exhibit 1: The reduction in per episode payments was statistically significant during the first two years of the model; after accounting for reconciliation payments, Medicare likely achieved savings



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

Notes: Ranges based on 90% CI are plotted as gray bars for reductions in non-standardized paid amounts and Medicare savings. Reductions in non-standardized paid amounts are based on estimates from a difference-in-differences model of per-episode standardized paid amounts that have been multiplied by negative one and converted to non-standardized amounts.

Negative savings reflect Medicare losses.

Reconciliation payments do not take on a range of values because the values are not based on estimates from a statistical regression.

CI=confidence interval.

3. *What types of CJR participant hospitals did and did not receive reconciliation payments?*

- **CJR participant hospitals that received reconciliation payments differed from those that did not.** The majority of CJR participant hospitals received reconciliation payments in at least one of the first two performance years; approximately one-quarter did not receive a reconciliation payment in either year. CJR participant hospitals with various characteristics that we examined received reconciliation payments, even though the CJR model included a broad range of hospitals due to its mandatory, random design. CJR participant hospitals that received a reconciliation payment in both performance years tended to have more LEJR episodes, start the CJR model with episode costs below their quality-adjusted target price, have higher quality scores, and have lower average patient complexity than hospitals that did not receive reconciliation payments in both years. CJR participant hospitals that did not receive a reconciliation payment in at least one year had lower LEJR volume,

performed a smaller share of the LEJRs in their markets, and had higher average patient complexity than hospitals that received reconciliation payments. Hospitals that received reconciliation payments for the first time in PY2 had a less complex patient mix in PY2 than they did in PY1.

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

- **Decreases in average episode payments were driven by reductions in the use of institutional PAC.** CJR participant hospitals discharged relatively fewer patients to an IRF (a 27.4% decrease from the CJR baseline proportion, $p<0.01$), resulting in a relative decrease in IRF payments of $357 for CJR episodes ($p<0.01$). There was a relative decrease of $508 in SNF payments ($p<0.01$), driven by a 2.3 day relative decrease in the number of days of SNF care ($p<0.01$). More CJR patients were first discharged to an HHA, although there was no change in the proportion of patients that received HHA services during the entire episode; thus, there was no statistically significant change in HHA payments.

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

- **Quality of care was maintained under the CJR model.** We observed no statistically significant changes in the quality of care for CJR episodes relative to control group episodes, as measured by the readmission rate, emergency department visits, and mortality.

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

- **By the end of the episode, CJR and control survey respondents reported making similar gains in functional status from before their hospitalization.** Self-reported pain did not differ between the two groups. For those patients discharged from the hospital to an IRF, SNF, or HHA, a smaller proportion of CJR patients improved their functional status than control patients while in the PAC setting. CJR and control respondents reported similar satisfaction with their overall recovery and care management and had similar care transitions experiences, although CJR respondents reported more reliance on caregiver help during their recovery.

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

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

- **The CJR model had no statistically significant impact on the volume of elective LEJR discharges.** The estimated impact of the CJR model on market-level LEJR discharge rates was a decrease of 0.033 discharges per 1,000 FFS beneficiaries, an estimate that is small and not statistically significant.

b. Are there any indications that the CJR patient population was healthier in the intervention period than in the baseline period?

- **CJR patients with elective LEJRs, and without major complications or comorbidities, the least complex and largest episode category, were healthier in the intervention period, on average, relative to control patients.** At the same time, there was no increase in the volume of LEJR discharges. There was also no change in the ratio of elective episodes without major complications or comorbidities to those with major complications or comorbidities. Taken together, these findings suggest that within the least complex elective episode category, there was a decrease in complex patients receiving LEJR and an increase in healthier patients receiving the surgery at CJR participant hospitals. This indicates that at least a portion of the decrease in episode payments may be due to a healthier mix of patients at CJR participant hospitals rather than care redesign.

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

- **The CJR model had no statistically significant impact on payments for services provided during the 30 days following the episode.** This indicates that CJR participant hospitals did not shift services until after the end of the 90-day post-hospital discharge period to lower average episode payments.

8. How did the impact of the CJR model vary for particular subgroups of MSAs, hospitals, and beneficiaries?

- **The CJR model resulted in statistically significant decreases in average episode payments for all subgroups of MSAs, hospitals, and episodes that we examined.** Average episode payments decreased in high-payment and low-payment MSAs, and the reductions were not statistically different from one another. Episode payments also decreased for hospital groups defined by their volume of LEJR episodes. Furthermore, we observed reductions in episode payments for both fracture and elective episodes. Finally, while average episode payments decreased for all patient complexity subgroups, episode payments decreased more for the more complex episodes than for the least complex episodes.
- **For fracture episodes, shifts from more intensive to less intensive PAC settings and a reduction in readmission payments drove the decrease in average episode payments.** CJR participant hospitals discharged relatively fewer patients with fracture episodes to an IRF (a 16.6% decrease from the CJR baseline proportion, $p<0.01$) and relatively more to a SNF (3.7% increase from the CJR baseline proportion, $p<0.05$) or HHA (16.5% increase from the CJR baseline proportion, $p<0.05$). This resulted in a decrease in average IRF payments per episode ($625, $p<0.01$) relative to the control group. This, along with a relative decrease in average readmission payments per fracture episode ($130, $p<0.10$), drove the reduction in average fracture episode payments ($1,267, $p<0.01$).

9. What key factors and model features influenced hospitals' choice of response to the CJR model?

- **Generally, hospital representatives we interviewed indicated they made decisions about their response to the model in the context of the hospital's market, complete orthopedic service line, internal resources, and past experiences.** Many interviewees indicated that the opportunity to prepare for future bundled payment models was a strong motivating factor in hospitals' response to the model. Hospital representatives indicated that prior hospital initiatives or participation in other payment and delivery models helped prepare for the CJR model. Hospital respondents often reported that the influence of the CJR model was not distinguishable from market factors that influenced decisions about the orthopedic service line.

10. What did CJR participant hospitals do to redesign care for their LEJR patients?

- **Hospital representatives reported initiating discharge planning before the hospital admission to educate patients that the goal is discharge directly home (with home health or outpatient therapy), and to identify high-risk patients to optimize health outcomes.** To reduce length of stay, hospitals implemented changes to pain management and physical therapy services. Hospital representatives also reported extending patient follow-up for a longer period and developing PAC protocols and preferred PAC provider networks to strengthen post-discharge outcomes. Some hospital representatives reported sharing internal cost savings with partnering providers, such as orthopedic surgeons, which is allowed under the CJR model through the gainsharing flexibilities, to reward efforts to reduce internal hospital costs.

11. How were relationships with orthopedic surgeons and PAC providers impacted by the CJR model?

- **Hospital representatives said that surgeon engagement in care redesign activities and communication between hospital and PAC staff improved under the CJR model.** Hospital interviewees described engaging surgeons in efforts to redirect patient discharge destination from SNFs to home and improve care coordination after discharge to reduce readmissions. Hospital interviewees that participated in gainsharing or that shared CJR performance data used these strategies for engaging physicians in their hospital's activities related to the model. PAC provider interviewees reported increased collaboration with other health care providers, including hospitals, orthopedic surgeons, primary care providers, and other PAC providers regarding LEJR patients.

C. Discussion

This second annual evaluation report demonstrates that the CJR model continues to be a promising approach to reducing payments for an episode of care that begins with LEJR surgery. This evaluation indicates that a range of hospitals, with varying resources and circumstances, can and do respond to the incentives under a mandatory episode-based payment approach for LEJR episodes to reduce per episode payments while maintaining quality. In response to the CJR model, participant hospitals said they continued with care redesign and engaged in strategies to discharge patients to the most appropriate setting after hospital discharge. These actions are consistent with the goals of the CJR model to improve care coordination by encouraging hospitals to work with physicians and PAC providers to be accountable for the entire episode. Even with reductions in the use of institutional PAC, quality of care was maintained.

While average episode payments for an LEJR decreased under the model, these lower episode payments result in savings to the Medicare program only if the aggregate reduction was greater than the reconciliation payments paid to participant hospitals. After accounting for the reconciliation payments, as well as the confidence interval around the estimated decrease in episode payments and the volume of CJR episodes, we estimate that the CJR model likely resulted in savings of $17.4 million to Medicare. However, there was a wide range around the estimate with losses of up to $41.2 million to savings of up to $75.9 million. While this may lead some to conclude that Medicare could lower quality-adjusted target prices to ensure Medicare savings, we do not know if participant hospitals would have made similar decisions about how to respond or reduced episode payments under different model design specifications.

In future reports we will continue to expand our understanding of the payment decreases under the CJR model and whether the lower episode payments translate into savings for the Medicare program. With additional time under the model, we also will have more information to evaluate the impact of changes in PAC on longer term patient outcomes. We also note that beginning in PY3, there were significant changes to the CJR model. Hospitals in the lower payment MSAs are no longer required to participate in the CJR model and were given a one-time opportunity to voluntarily continue. The next annual report will examine the impact of the CJR model on the hospitals that remained mandatory participants and on all hospitals that ever participated in the CJR model.

I. Introduction

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

The second annual CJR model evaluation report presents findings from the first two performance years, which include episodes initiated on or after April 1, 2016 that ended by December 31, 2017. During this period, the model was mandatory for nearly all hospitals in 67 geographic areas, defined by MSAs. At the start of PY3, the number of mandatory MSAs was scaled back to the 34 MSAs with the highest historical payments; hospitals in the other 33 MSAs were given the opportunity to continue to participate voluntarily.

This report provides results from the largest expression of the CJR model and encompasses the broadest range of CJR model participant hospitals.

A. The CJR Model

CJR participant hospitals are financially accountable for the cost and quality of health care services during an LEJR episode of care. They have incentives to reduce episode payments and improve quality by coordinating care with the physicians, PAC providers, and other providers and clinicians involved in the episode.[9] At the end of each performance year, payments for the episodes initiated at a participant hospital are compared to the hospital's quality-adjusted target price to determine whether the hospital earns a reconciliation payment or repays Medicare based on its performance. The CJR model has a mandatory, randomized design in which hospitals in randomly selected MSAs were required to participate. Because of this design, a spectrum of hospitals with varying levels of infrastructure, care redesign experience, episode costs and utilization, and market positions are participating, which allows a broad test of the CJR model.

Episode definition. Under the CJR model, an LEJR episode of care begins with the hospitalization for the surgery and extends through the 90 days after hospital discharge. All Medicare-covered items and services provided during this period, with some exclusions, are included in the episode bundle.[10] All providers and suppliers involved in the episode continue to be paid under Medicare's FFS payment systems.

---

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

10 Excluded items, services, and payments include: hemophilia clotting factors; new technology add-on payments; transitional pass-through payments for medical devices; items and services unrelated to the anchor hospitalization as specified by CMS on the CJR model website, including (i) inpatient hospital admissions for MS–DRGs for oncology, trauma medical, chronic disease surgical, and acute disease surgical diagnoses, (ii) Medicare Part B services for acute disease and certain chronic disease diagnoses, (iii) certain per beneficiary per month payments; certain incentive programs and add-on payments under existing Medicare payment

Annual reconciliation. After the end of each model performance year, CMS reconciles the payments for each participant hospital's LEJR episodes against the hospital's quality-adjusted target price. The quality-adjusted target price is based on a discounted blend of the hospital's average historical episode payments and the region's average historical episode payments. During the first two performance years, two-thirds of the quality-adjusted target price is the hospital's average historical episode amount and one-third is the regional average. By PY4 the quality-adjusted target price is based entirely on the historical regional amount, which is intended to reduce variation in LEJR episode payments and reward hospitals for reducing payments below their regional peers. The discount to the quality-adjusted target price is intended to be Medicare's portion of the decrease in spending under the model. At reconciliation, the discount is adjusted based on the participant hospital's composite quality score. A lower discount is applied to the target price for participant hospitals with a higher quality score (resulting in a higher quality-adjusted target price) to encourage participant hospitals to focus on improving quality.

Hospitals with LEJR episode payments *below* their quality-adjusted target price and an "acceptable" or higher composite quality score earn a reconciliation payment. The reconciliation payment equals the difference between the quality-adjusted target price and actual episode spending, up to a stop-gain limit. Starting in PY2, hospitals with episode payments *above* their quality-adjusted target price repay Medicare a portion of the difference, subject to a stop-loss limit. In PY1, this repayment responsibility was forgiven to allow hospitals time to gain experience under the CJR model before implementation of two-sided risk.

To account for the variation in the complexity and resulting costs of LEJR episodes, the quality-adjusted target price is risk-adjusted based on the presence or absence of major complications and comorbidities and presence of fracture. This risk adjustment approach results in four quality-adjusted target prices for each hospital. The purpose of the risk adjustment is to reduce any unintended incentive to avoid patients with higher costs due to their greater needs.

Mandatory, randomized design. The mandatory, randomized design of the CJR model ensures a valid control group for assessing the model's impact. For the two performance years that are covered in this annual report, all acute care hospitals paid under the Medicare IPPS, with few exceptions, in 67 randomly selected MSAs were required to participate. The mandatory MSAs were identified from 171 MSAs that were eligible for participation when the model design was finalized. MSAs were selected for participation using eight sampling strata based on a median split of MSA population size and quartiles of average MSA historical episode payments.[11] An MSA's probability of selection increased with the payment quartiles in order to oversample high payment MSAs for participation in CJR. This was because of the belief that there is greater need and more

systems; and payments for otherwise included items and services in excess of two standard deviations above the mean regional episode payment.

[11] Originally, 196 MSAs were identified as eligible for participation in the CJR model and the mandatory MSAs were randomly selected from this pool. CMS later identified 25 MSAs that were ineligible for selection after accounting for BPCI PGP participation.

opportunities for payment reductions in higher payment areas. Eligible MSAs that were not selected are a natural control group for evaluating the impact of the CJR model.

The model's design also results in a diverse group of CJR participant hospitals, including hospitals that might not voluntarily participate in an episode-based payment model. CJR participant hospitals were distributed across the country, with representation from high and low cost markets. They also varied in terms of their ownership, teaching hospital status, LEJR volume, and size, among other factors (Appendix C).

CJR participant hospitals were generally similar to all other hospitals paid under Medicare's IPPS, with few exceptions. Because the mandatory participation is based on the hospital's MSA, all CJR participant hospitals were located in urban areas, compared with almost 94% of all other hospitals paid under Medicare's IPPS ($p<0.01$).[12] Other differences between the CJR participant hospitals and all other IPPS hospitals reflect this urban focus. CJR participant hospitals included a higher proportion of teaching hospitals than the universe of other IPPS hospitals (42% vs. 34%, $p<0.01$). CJR participant hospitals were more likely to be safety-net hospitals (40% vs. 29%, $p<0.01$). They were also larger than the typical IPPS hospital, with an average of 253 beds, compared with 196 beds ($p<0.01$). CJR participant hospitals were less likely to have had prior experience with the Bundled Payments for Care Improvement initiative (10% vs. 18%, $p<0.01$), again owing to the MSA selection approach in which CMS chose to sample from MSAs with low Bundled Payments for Care Improvement initiative penetration.

By design, CJR participant hospitals had higher average historical episode payments. They had average episode payments of $30,546, compared with $28,700 for all other IPPS hospitals ($p<0.01$). A larger percentage of episodes initiated in CJR participant hospitals were first discharged to a SNF (46% vs. 43%, $p<0.01$) and a smaller percentage were first discharged home without home health (HH) (13% vs. 16%, $p<0.01$) (Appendix C).

For more information about the CJR model, please visit https://innovation.cms.gov/initiatives/cjr and reference the first annual evaluation report.

B. Evaluation Conceptual Framework

The conceptual framework for the evaluation of the CJR model (Exhibit 2) reflects the fundamental features of the model and is informed by health services research literature, including evaluations of other bundled payment approaches.[13] The evaluation framework focuses on the hospital where the LEJR episode begins because the hospital has the incentives to control payments and improve quality across the entire episode. The hospital's decisions about whether and how to respond to the model will reflect its resources and market conditions. The impact of the

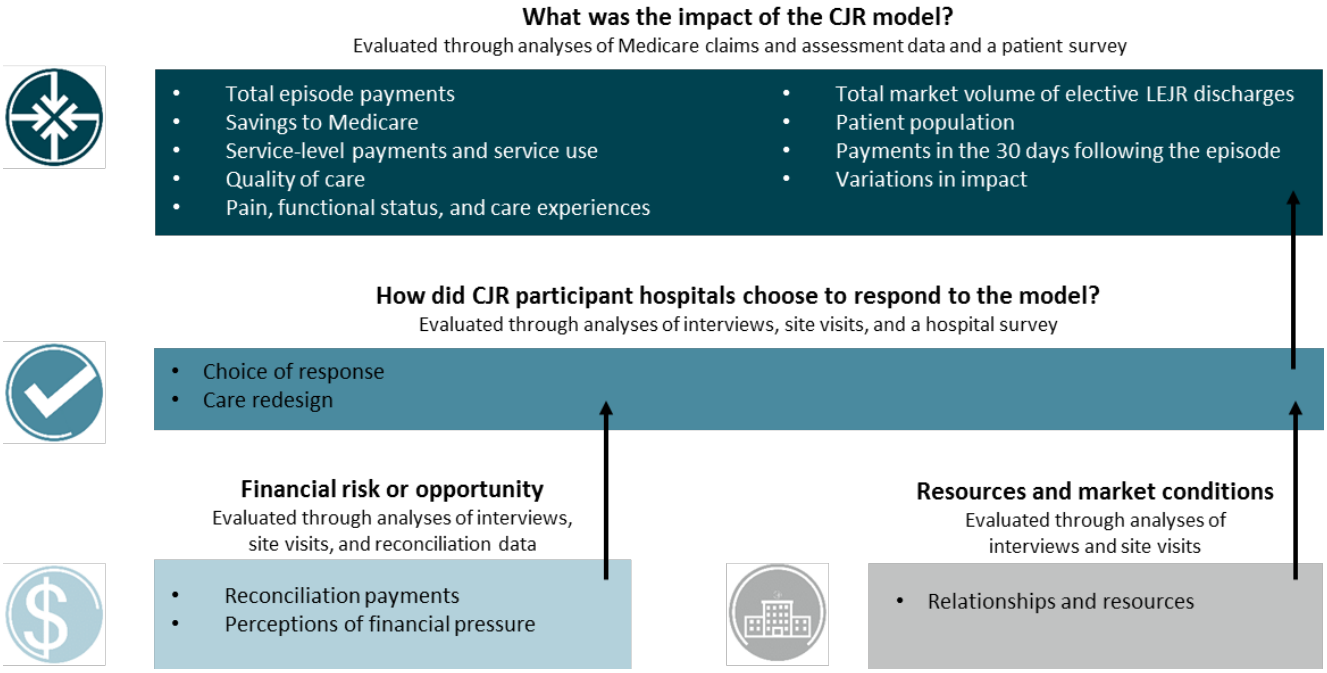
[12] Appendix C – CJR vs. all other IPPS Hospitals; Appendix D – Definitions of Hospital and Patient Characteristics.

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

CJR model will be influenced by those decisions, as well as whether the hospital's episode payments are above or below its quality-adjusted target price.

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

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



Note: LEJR = lower extremity joint replacement.

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

Medicare claims reveal impacts of the CJR model on quality outcomes. Patient assessment data and self-reported measures from a patient survey provide information on functional status and pain. The cross-sectional patient survey analysis compares patients in CJR episodes with patients in

control episodes, providing insights into the relationship between CJR participation and patient experience.

The claims analysis reveals whether the CJR model resulted in participant hospitals reducing episode payments as intended, but additional analysis is needed to determine if participant hospitals responded to the model by increasing the volume of episodes, which would raise total Medicare spending, rather than reduce Medicare spending. We examine whether the change in volume of LEJR hospital discharges differs between mandatory CJR MSAs and control MSAs.

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

Choice of response. Hospital leaders must consider multiple organizational factors, in addition to the potential for financial risk or opportunity, and internal and external resources, in making the business case for whether and how to respond to the CJR model. Orthopedic surgery is one of multiple service lines that compete for staff and other resources. The CJR model is one initiative that may or may not align with initiatives from other payers, state-specific policies, local labor markets, and other factors. Site visits, structured interviews, and hospital surveys provide data about how hospitals made decisions about their response to the model. They also reveal how hospital administrators use internal data or the claims-based episode data provided by CMS to determine the actions they will take to affect episode spending and quality. (Section II.J examines the actions participant hospitals reported taking in response to the model.)

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

Resources and market conditions. A hospital's internal resources and market conditions will provide opportunities or constraints on its responses to the model. Hospitals with more capital and operational resources, such as dedicated care coordination staff or robust health information technology infrastructure, may be better situated to redesign care for LEJR episodes. Other hospital resources – such as leadership support, experience

with bundled payments or similar payment models, or ownership of PAC providers or employment of surgeons – may also affect their choices as well as their success in reducing payments below their quality-adjusted target price. Market conditions, such as the supply and characteristics of other providers involved in the episode, will affect how and whether hospitals garner support for delivering care more efficiently during the episode. Hospital representative interviews provide information about how they perceive the impacts of the actions they have implemented. The site visits provide rich information about how market conditions and particular hospital resources affect responses to the model. (Section II.I examines the factors that influenced participant hospitals' response to the model; detailed case studies can be found in the Supplement to this report.)

II. Results

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

The CJR model was designed to reduce average payments for an episode of care for LEJR. While participant hospitals are held accountable for payments for the entire episode, all providers and suppliers involved in the episode continue to be paid under Medicare's FFS payment systems. Under Medicare's FFS payment systems, providers are paid based on the volume of services rather than the value or quality of services; thus, FFS payment systems may result in fragmented, unnecessary, or duplicative care.[14] By holding participant hospitals accountable for the payments and quality for an episode of care, the CJR model aims to encourage hospitals to move towards value-based care.

1. Key findings

- Average standardized allowed amounts (payments) decreased by $997 more for CJR episodes than for control group episodes during the first two performance years of the CJR model. This equates to a 3.7% decrease from the baseline.

2. Methods

The analysis uses a difference-in-differences (DiD) design to estimate the differential change in average Medicare standardized allowed amounts (payments) between the baseline (April 2012 through March 2015) and intervention period (April 2016 through December 2017) for beneficiaries who received LEJRs from CJR participant hospitals relative to beneficiaries who received LEJRs from control group hospitals. We use standardized payments to ensure that observed payment differences reflect actual differences in billed services rather than Medicare payment policies. We use allowed amounts to eliminate variation in payments that could stem from whether beneficiaries have met their deductible when they had the LEJR surgery. We use the DiD method because it controls for common trends and fixed differences in outcomes that may occur between CJR hospitals and the control group hospitals. In addition, we control for beneficiary, market, and hospital characteristics that can vary over time and between the CJR and control group. The control group MSAs are weighted to be representative of the distribution of the CJR MSAs across the eight sampling strata. The percent decrease in payments represents the percent change from the CJR baseline that is due to the CJR model. It is calculated by dividing the DiD

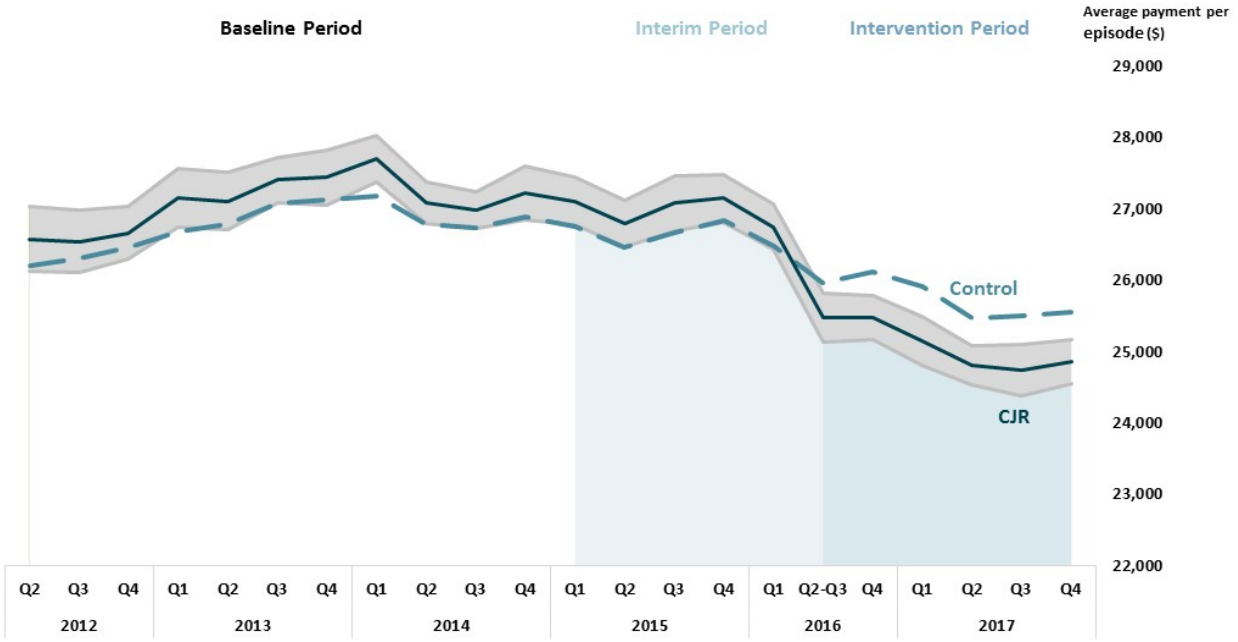
Standardized allowed amounts are used in this report to measure the impact of the CJR model on average episode and service-level payments. *Standardization* removes wage adjustments and other Medicare payment adjustments. *Allowed amounts* include beneficiary cost sharing.

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

estimate by the CJR baseline average. Additional details about the methodology are in Appendix E.

3. Results

Exhibit 3: Average episode payments declined more for CJR episodes than for control episodes in PY1-2



Source: Lewin analysis of Medicare claims and enrollment data for episodes initiated on or after January 2012 that ended by December 2017.

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

The gray shading represents the 95% confidence interval for the CJR estimate.

PY = performance year.

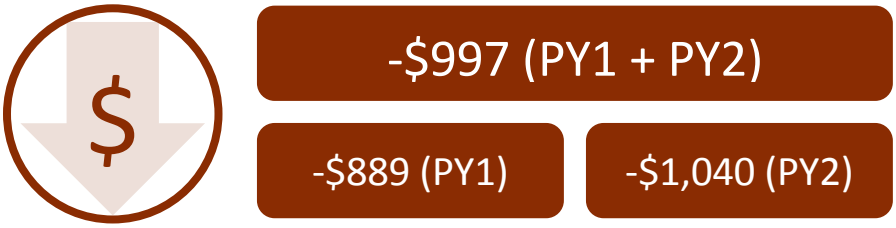
During the first two performance years, the CJR model resulted in a relative reduction in average payments for an LEJR episode of care. This analysis of average episode payments does not incorporate reconciliation payments made to CJR participant hospitals; therefore, the results do not represent savings to the Medicare program. An analysis of Medicare savings is in Section II.B.

While average episode payments declined for both CJR and control group episodes during the first two performance years of the CJR model, payments declined more for CJR episodes (Exhibit 3). Average episode payments decreased by $997 more for CJR episodes than for control group episodes from the baseline to the intervention period ($p<0.01$, Exhibit 3). This relative reduction equates to a 3.7% decrease in average episode payments for CJR episodes from the baseline.

Average episode payments decreased in both performance years under the model. Average episode payments decreased by $889 more for CJR episodes than for control episodes in the first

performance year ($p<0.01$) and by $1,040 more for CJR episodes in the second performance year ($p<0.01$, Exhibit 4).

Exhibit 4: Average CJR episode payments decreased in performance years 1 and 2



Source: Lewin analysis of Medicare claims and enrollment data for episodes initiated in 2012 through 2014 that ended between April 2012 and March 2015 (baseline) and episodes initiated during or after April 2016 that ended by December 2017 (intervention).

Note: PY = performance year.

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

Medicare savings due to the CJR model reflects the change in average episode payments, the reconciliation payments made to or received from hospitals under the model, and any changes in the volume of LEJR as a result of the model. As detailed in Section II.A, CJR participant hospitals reduced average standardized episode payments for LEJR episodes during the first two performance years. As detailed in Section II.G.1, the CJR model did not have a statistically significant impact on the volume of elective LEJR discharges in mandatory CJR markets. This section presents estimated Medicare savings based on the change in average non-standardized paid amounts and reconciliation payments made to or received from CJR participant hospitals.

1. Key findings

- Estimated total Medicare savings from the CJR model was $17.4 million, ranging between Medicare losses of $41.2 million to Medicare savings of $75.9 million.
- Estimated Medicare savings was $117 per episode, ranging from a loss of $278 per episode to a savings of $513 per episode (based on a 90% confidence interval).

2. Methods

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

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

Reconciliation payments are the payments made to CJR participant hospitals by Medicare for meeting cost and quality targets and repayments from CJR participant hospitals to Medicare for

failing to meet cost and quality targets.[15] To calculate Medicare savings we use non-standardized paid amounts instead of standardized allowed amounts that we use in other sections. We use non-standardized paid amounts for this analysis because they are the actual payments made from Medicare to providers incorporating geographic and other payment adjustments and excluding beneficiary cost sharing. See Appendix E for additional details about these methods.

3. Results

During the first two performance years, the CJR model likely reduced Medicare program spending by an estimated $17.4 million. Accounting for the uncertainty in per episode savings, however, results in estimated savings due to the CJR model that ranges from Medicare losses of $41.2 million to Medicare savings of $75.9 million (Exhibit 5). Medicare savings is based on an estimated reduction in per episode non-standardized paid amounts of $989, less $872 in average reconciliation payments per episode. This results in an estimated savings of $117 per episode, ranging from an increase to Medicare of $278 to a savings of $513 per episode, accounting for uncertainty through a 90% confidence interval. Multiplying the per-episode savings estimates by 147,923 episodes from CJR participant hospitals yields the total Medicare savings amounts.

Exhibit 5: The CJR model likely saved Medicare money, but there is a wide range around savings estimates

	Estimate	Range (based on 90% CI)
Reduction in non-standardized paid amounts per episode	**$989**	$593 to $1,385
Average reconciliation payment per episode	**$872**	
Medicare savings per episode	**$117**	–$278 to $513
Number of episodes	**147,923**	
Total Medicare savings	**$17,362,296**	–$41,163,818 to $75,888,410

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

Notes: Reductions in non-standardized paid amounts are based on estimates from a difference-in-differences (DiD) model of per-episode standardized paid amounts that have been multiplied by negative one and converted to non-standardized amounts.

Negative savings reflect Medicare losses.

CI=confidence interval.

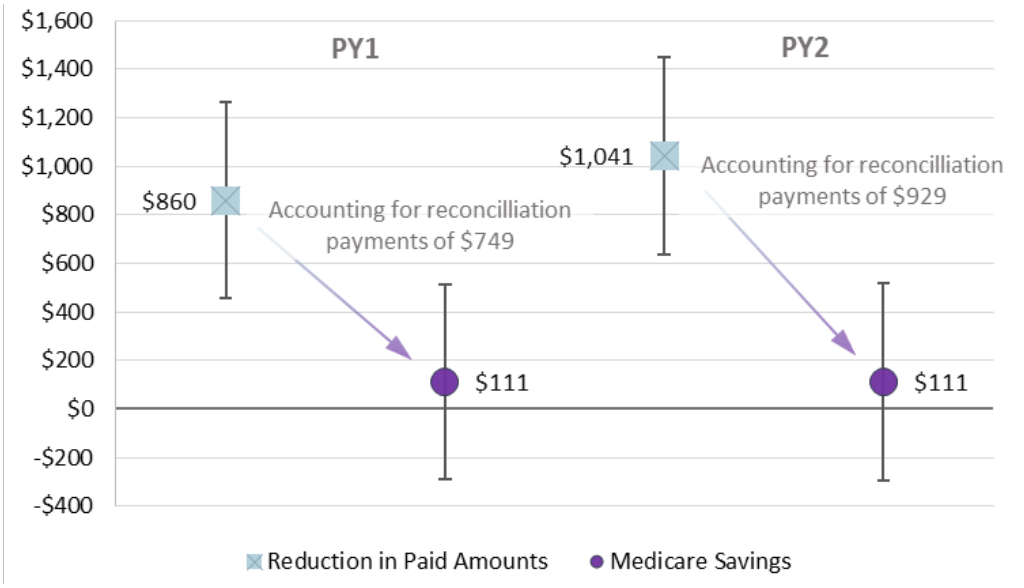
We also estimated Medicare savings separately for each performance year. Average non-standardized paid amounts per episode decreased by $860 and $1,041, for PY1 and PY2 respectively. The range of per-episode reductions for each performance year, based on 90% confidence intervals, overlapped considerably (Exhibit 6).[16] After accounting for reconciliation

[15] Reconciliation payments can be positive or negative, depending if the payment is from Medicare to a participant or a repayment from the participant to Medicare. In the program literature they are often referred to by the technical term "net payment reconciliation amounts" or "NPRA."

[16] These calculations are converted impact estimates that show for each performance year a relative reduction in Medicare standardized payments for an LEJR episode of care. The DiD estimate for standardized paid amounts

payments ($749 in PY1 and $929 in PY2), we estimated Medicare savings of $111 per episode in both PY1 and PY2.

Exhibit 6: The reduction in per episode payments was statistically significant in CJR performance years 1 and 2; after accounting for reconciliation payments, Medicare likely achieved net savings



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

Notes: Ranges based on 90% CI are plotted as gray bars for reductions in non-standardized paid amounts and Medicare savings. Reductions in non-standardized paid amounts are based on estimates from a difference-in-differences (DiD) model of per-episode standardized paid amounts that have been multiplied by negative one and converted to non-standardized amounts.

Negative savings reflect Medicare losses.

Reconciliation payments do not take on a range of values, because the values are not based on estimates from a statistical regression.

CI=confidence interval, PY = performance year.

The CJR model did not incorporate downside risk for the first performance year, which means that participant hospitals with average episode payments above their quality-adjusted target price were not responsible for repayments. Had participant hospitals been required to make repayments in PY1, reconciliation payments per episode would have been reduced from $749 to $525, resulting in estimated Medicare savings of $335 per episode. Medicare savings from the first two performance years would have been $189 per episode and $28.0 million in total, ranging between total Medicare losses of $30.5 million to total Medicare savings of $86.5 million (accounting for uncertainty from estimated components based on a 90% confidence interval). Thus, had downside

for PY1 was –$833 ($p<0.01$) with a 90% confidence interval ranging from –$1,223 to –$443, and the DiD estimate for PY2 was –$1,006 ($p<0.01$) with a 90% confidence interval ranging from –$1,399 to –$612.

risk not been waived in PY1, our estimate of Medicare savings would have been higher, however, we still would not be able to conclude with certainty that the CJR model resulted in savings.

The results presented in Exhibit 5 and 6 use confidence interval approaches to inform us about the statistical uncertainty in our estimates. To express our findings in an alternative way, we also calculated the probability that the CJR model resulted in Medicare savings.[17] There was a 69% probability that the CJR model resulted in Medicare savings greater than zero for the first two performance years. Had participant hospitals been required to make repayments in PY1, the probability of savings greater than zero would have increased to 79%.

C. What types of CJR participant hospitals did and did not receive reconciliation payments?

At the end of each performance year, Medicare compares episode payments for a CJR participant hospital to its quality-adjusted target price to determine whether the hospital should receive a reconciliation payment (beginning in PY1) or make a repayment to Medicare (beginning in PY2). The underlying assumption of the CJR model is that the opportunity to receive reconciliation payments will serve as an incentive for CJR participant hospitals to invest in care redesign and coordination efforts, with the goal of increasing the efficiency and quality of care provided to patients undergoing LEJR. As discussed in Section II.A above, average episode payments decreased under the CJR model, suggesting increased efficiency in care delivery for CJR patients. This section focuses on the characteristics of hospitals that did and did not receive reconciliation payments.

The characteristics of hospitals that did and did not receive reconciliation payments is particularly of interest for this model. In voluntary models, it is assumed that hospitals elect to participate, in part, based on their expectations of how likely they are to achieve reconciliation payments. As a result, participant hospitals are unlikely to be representative of all hospitals. Because the CJR model is mandatory, the examination of the performance of its hospitals provides an opportunity to explore the response of a broad range of hospitals—with different infrastructures, care redesign experiences, patient populations, utilization patterns, and market positions—to episode-based payments. This includes hospitals that would not have elected to participate in other circumstances.

[17] Specifically, we calculated the probability of the value of the impact estimate required to achieve Medicare savings greater than or equal to zero. Appendix E explains the calculation of the probabilistic statements in more detail.

1. Key findings

- A large majority of CJR participant hospitals received a reconciliation payment: 44 % of CJR participant hospitals received reconciliation payments in both performance years and an additional 33% received a reconciliation payment in one of the two performance years; 23% never received reconciliation payments.
- While CJR participant hospitals with various characteristics received reconciliation payments, on average, hospitals that received reconciliation payments in both performance years had a higher volume of LEJR episodes, higher quality of care, and had lower average patient complexity than hospitals that never received reconciliation payments.

2. Methods

We compared groups of CJR participant hospitals based on whether they received reconciliation payments in both of the first two performance years, one performance year, or never received reconciliation payments. Hospitals that did not receive reconciliation payments included hospitals with episode payments above their quality-adjusted target prices and hospitals that were ineligible to receive reconciliation payments (i.e., hospitals that had episode payments below their quality-adjusted target price but failed to meet the minimum quality requirements).

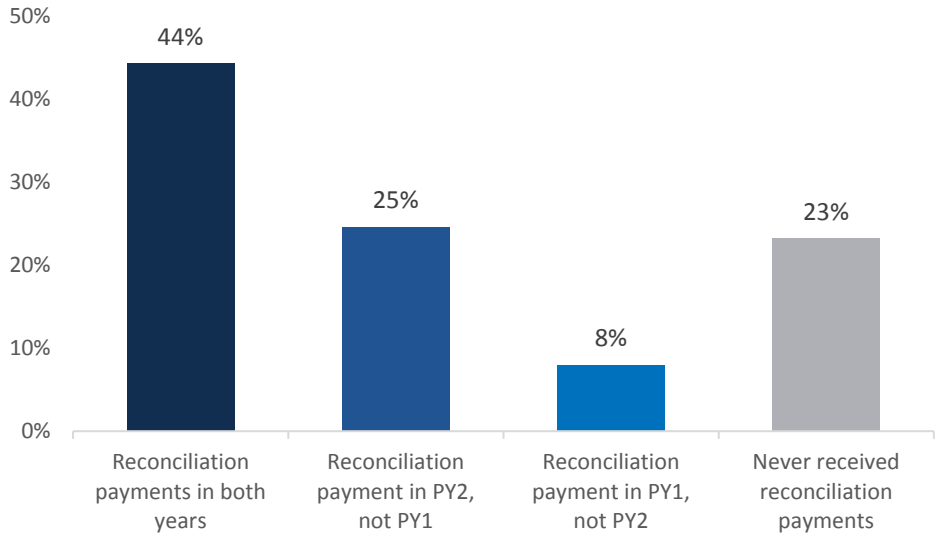
We tested for differences between groups using chi-square tests for categorical variables and analysis of variance for continuous variables. This analysis is descriptive, that is it explores the relationship between receiving reconciliation payments and one other variable at a time, and does not incorporate risk adjustment or other advanced statistical methods.

The analysis examines PY1 and PY2 reconciliation payments. The PY1 results are considered final as of the writing of this report. The results for PY2 are not final and need to be interpreted with caution because they are based on preliminary PY2 reconciliation data, which is subject to change. The final reconciliation occurs 14 months following the end of a performance year.

3. Results

In PY1, 52% of CJR participant hospitals received a reconciliation payment, and in PY2, 69% of CJR participant hospitals received a reconciliation payment. Across both performance years, 77% of CJR participant hospitals ever received a reconciliation payment with 44% receiving reconciliation payments in both performance years and 33% receiving a reconciliation payment in only one year (Exhibit 7). Less than one-quarter of CJR participant hospitals (23%) did not receive a reconciliation payment in either year.

Exhibit 7: 44% of hospitals received reconciliation payments in both performance years and 33% received a reconciliation payment in one of the two performance years under the CJR model



Source: Lewin analysis of CJR payment contractor data for CJR participating hospitals in PY1 (episodes starting on or after April 2016 and ending on or before December 2016) and PY2 (episodes ending between January and December 2017).

Notes: Hospitals that did not receive reconciliation payments include hospitals with average episode payments above their quality-adjusted target prices and hospitals that were ineligible to receive reconciliation payments.

PY = performance year.

CMS designed the CJR model to disproportionately represent markets (i.e., MSAs) and hospitals with the greatest opportunity for payment reductions by oversampling high-payment MSAs. Overall, 68% of hospitals were in high-payment MSAs. Hospitals that received reconciliation payments in both years were less likely to be located in high-payment MSAs. Sixty-two percent of hospitals that received reconciliation payments in both performance years were located in high-payment MSAs, compared with 77% of hospitals that never received reconciliation payments (Exhibit 8, $p<0.05$). Although hospitals in high-payment MSAs reduced episode payments (Section II.H.3.a), the reductions were not always enough to come under the quality-adjusted target price.

Further, hospitals that received reconciliation payments in both years were more likely to have started the CJR model with average historical payments below their quality-adjusted target price (45% compared with 24% of hospitals that never received reconciliation payments, $p<0.01$). This may imply that some CJR participant hospitals receive reconciliation payments because they already had low episode payments.

The CJR model also rewards hospitals with higher quality, as measured by a composite quality score, by reducing the effective discount percentage applied at reconciliation. As expected, hospitals that received reconciliation payments in both performance years had higher composite quality scores, averaging 12.4 in both years (out of a total of 20 points), while average composite

quality scores were approximately 5 points lower for hospitals that never received reconciliation payments (7.2 in PY1 and 7.0 in PY2, p<0.01).

Exhibit 8: Hospitals that received reconciliation payments in both performance years differed from hospitals that did not

	Reconciliation payments in both years	Reconciliation payments in PY2, not PY1	Reconciliation payments in PY1, not PY2	Never received reconciliation payments
NUMBER OF HOSPITALS	300	167	54	157
MARKET				
Percent located in a high payment market*	62%	67%	72%	77%
Average LEJR market share**	13%	8%	5%	7%
HOSPITAL				
Average number of episodes in first two years**	307	174	99	126
Average percent of discharges that were LEJR**	9%	6%	4%	4%
Percent that started CJR below their target price**	45%	33%	22%	24%
Average composite quality score				
Year 1**	12.4	9.6	9.7	7.2
Year 2**	12.4	10.6	7.5	7.0
PATIENT				
Average patient complexity score				
Year 1**	1.06	1.30	1.27	1.38
Year 2**	1.06	1.18	1.38	1.40
Average percent with fracture				
Year 1**	15%	24%	32%	30%
Year 2**	16%	23%	37%	29%

Source: Lewin analysis of CMS payment contractor data and CJR quality performance data for CJR participating hospitals in PY1 (episodes starting on or after April 2016 and ending on or before December 2016) and PY2 (episodes ending between January and December 2017), and Medicare claims data for episodes initiated in 2012 through 2014 that ended between April 2012 and March 2015 (baseline) and episodes initiated on or after April 2016 that ended by December 2017 (intervention).

Notes: Hospitals that did not receive reconciliation payments include hospitals with average episode payments above their quality-adjusted target prices and hospitals that were ineligible to receive reconciliation payments. Patient complexity is measured by hierarchical condition category (HCC) score with scores of greater than 1.00 indicating higher patient complexity.

* Significant difference at the p<0.05 level. ** Significant difference at the p<0.01 level.

LEJR = lower extremity joint replacement, PY = performance year.

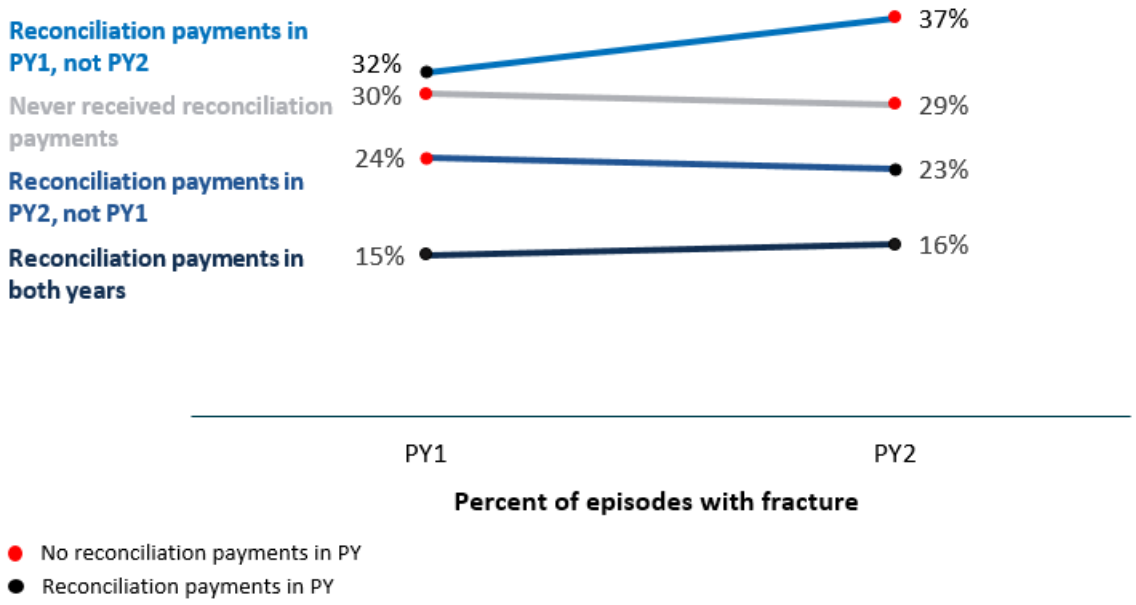
Hospitals that received reconciliation payments in both years differed from hospitals that never received reconciliation payments on other factors. They were more likely to be not-for-profit (68% vs. 50%, p<0.01) and have a lower average proportion of low-income patients (average disproportionate share hospital (DSH) patient percentage of 26% vs. 37%, p<0.01).

There is a relationship between the relative size of the hospital's LEJR service line and performance under the CJR model (Exhibit 8). Hospitals that received reconciliation payments in both years had more LEJR episodes, on average (averaging 307 in the first two performance years vs. 126 for hospitals that never received reconciliation payments, p<0.01) and a higher proportion of LEJR discharges as a share of total discharges prior to the start of the model (9% vs. 4% for hospitals that never received reconciliation payments, p<0.01). They also had a higher proportion of the LEJR market share prior to the start of the model (they performed on average 13% of the

Medicare FFS LEJRs in their MSA vs. 7% for hospitals that never received reconciliation payments, p<0.01).

Hospitals that received reconciliation payments served LEJR patient populations with lower average complexity, defined by presence of a fracture or hierarchical condition category score, than those that did not. Compared to hospitals that received reconciliation payments in both years, hospitals that never received reconciliation payments had almost double the proportion of episodes with fractures (approximately 30% vs. 15%, p<0.01 for PY1 and PY2) (Exhibit 9). The higher proportion of episodes with fracture is likely correlated with hospital size, with smaller hospitals receiving a higher proportion of patients needing LEJR due to fracture.

Exhibit 9: Low proportion of episodes with fracture was associated with receiving reconciliation payments



Source: Lewin analysis of CMS payment contractor data for CJR participating hospitals in PY1 (episodes starting on or after April 2016 and ending on or before December 2016) and PY2 (episodes ending between January and December 2017), and Medicare claims data for episodes initiated in 2012 through 2014 that ended between April 2012 and March 2015 (baseline) and episodes initiated on or after April 2016 that ended by December 2017 (intervention).

Notes: Hospitals that did not receive reconciliation payments include hospitals with average episode payments above their quality-adjusted target prices and hospitals that were ineligible to receive reconciliation payments.

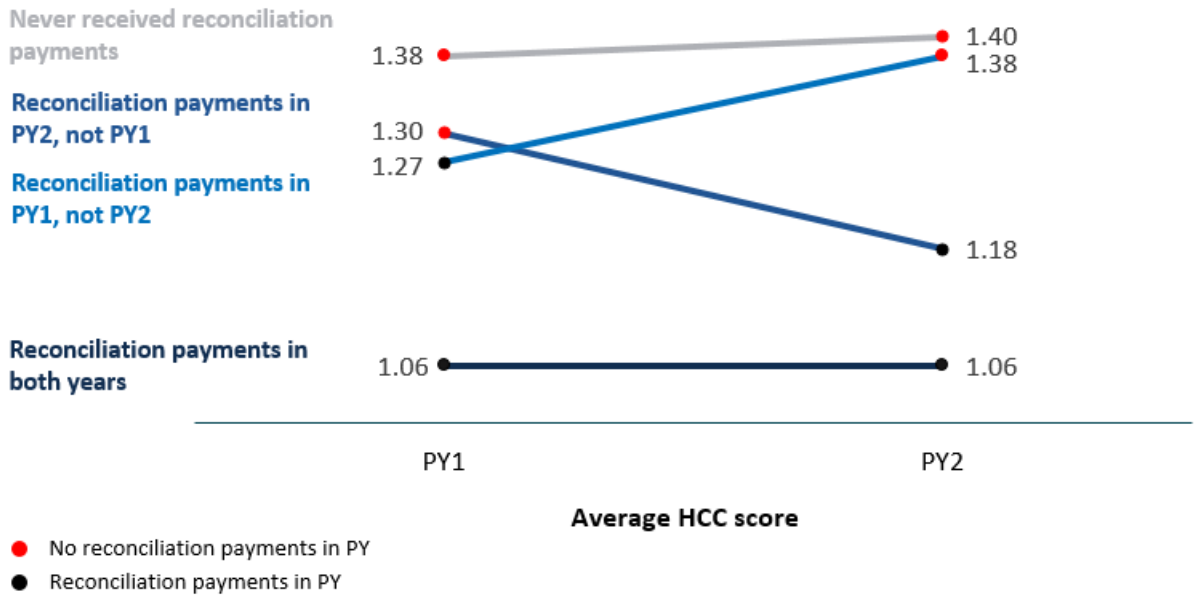
Fracture defined based on ICD codes for fracture provided by CMMI on the CJR website: https://innovation.cms.gov/Files/worksheets/cjr-icd10hipfracturecodes.xlsx.

PY = performance year.

Additionally, the average hierarchical condition category (HCC) score was correlated with receiving reconciliation payments, which suggests that hospitals serving LEJR patient populations with higher average complexity were less likely to receive a reconciliation payment (Exhibit 10). Hospitals that received reconciliation payments in both years had less complex patients on average, as indicated by lower average HCC scores for their LEJR episodes, than hospitals that never received reconciliation payments (1.06 vs. 1.38 in PY1, p<0.01; 1.06 vs. 1.40 in PY2, p<0.01).

Further, results suggest that hospitals that did not receive a reconciliation payment in PY1 but received a reconciliation payment in PY2 had a shift from higher patient complexity to lower patient complexity over this period ($p<0.05$, Exhibit 10). This may indicate that the simple risk stratification methodology used by CMS to set target prices was not sufficient to account for variations in patient complexity. (CMS uses a simple risk stratification methodology that assigns separate episode target prices for Medicare Severity-Diagnosis Related Groups (MS-DRGs) 469 and 470 and for patients with hip fractures within each MS-DRG.)

Exhibit 10: Lower average patient complexity was associated with receiving reconciliation payments



Source: Lewin analysis of CMS payment contractor data for CJR participating hospitals in PY1 (episodes starting on or after April 2016 and ending on or before December 2016) and PY2 (episodes ending between January and December 2017), and Medicare claims data for episodes initiated in 2012 through 2014 that ended between April 2012 and March 2015 (baseline) and episodes initiated on or after April 2016 that ended by December 2017 (intervention).

Notes: Hospitals that did not receive reconciliation payments include hospitals with average episode payments above their quality-adjusted target prices and hospitals that were ineligible to receive reconciliation payments.

Patients who are healthier than the average Medicare beneficiary will have HCC scores of less than 1.0, while patients who are unhealthier than the average Medicare beneficiary will have HCC scores of greater than 1.0.

HCC = hierarchical condition category, PY = performance year.

a. Conclusion

The financial performance of participating hospitals under the CJR model was associated with differences in institutional characteristics, a mix of episode types, LEJR volume, and patient complexity. While these correlations do not explain the underlying factors leading to the differences, they provide insights into the relationships that should be investigated further to understand how the impact of the CJR model varies across hospitals and episodes.

The majority of CJR participant hospitals received reconciliation payments in one or more performance years. While a range of hospitals with varied characteristics received reconciliation

payments under the CJR model, doing so appeared more challenging for some hospitals than others. Treating more complex episodes and patients, which may be beyond hospitals' control, were correlated with not earning a reconciliation payment, which suggests that the simple risk stratification methodology to set target prices may not adequately account for variations in patient complexity. Hospitals that received reconciliation payments in both years had higher LEJR volume, provided higher quality of care, treated LEJR patients with lower average complexity, and were more likely to be located in a historically low payment MSAs.

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

Changes in service-level payments and use provide insights into how hospitals reduced total episode payments. Shifts in patient mix across service settings provide additional indications of where and how hospitals focused their efforts to reduce episode payments. Whereas hospital payments are unlikely to change because Medicare payment is on a per-discharge basis, PAC payments, which comprise roughly one-third of total LEJR episode payments, can be reduced by shifting service use from more to less expensive care settings. In general, payments for IRF care are higher than payments for SNF care, and both of these institutional PAC settings tend to have higher Medicare payments than HHA services.

1. Key findings

- The relative decrease in average episode payments was driven by relative decreases in institutional PAC payments.
- A smaller proportion of CJR patients were discharged to an IRF and a larger proportion to HH than control patients.

2. Methods

This analysis uses a DiD design (described in Section II.A.2) to estimate the differential change in average standardized allowed amounts (payments) and average utilization by service during the 90 days following discharge from the hospital. Average payments by service are based on all episodes, including episodes that did not receive the particular service.

We also evaluated changes in the complexity of CJR patients discharged to IRF, SNF, or HHA relative to the control group. We used a similar DiD design to estimate the unadjusted differential change in patient complexity measures obtained from IRF, SNF, and HHA admission assessments, claims, and enrollment data. Assessment measures vary by discharge setting (Appendix G). For this analysis, the intervention period was six quarters (from April 2016 to September 2017), which was one-quarter shorter than the analyses that rely only on claims because of the longer runout time needed for the assessment data.[18]

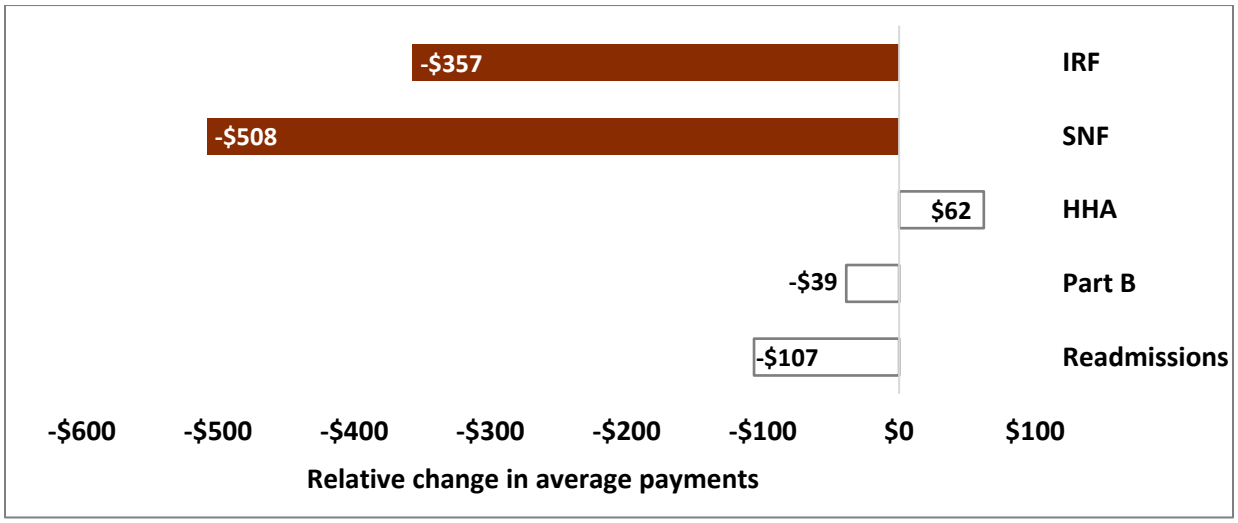
[18] While this intervention period is three months shorter than for other analyses in this report, this does not appear to affect our conclusions as the results have been stable over time.

3. Results

a. Service level payments and use

During the first two performance years, the relative decrease in average episode payments ($997, p<0.01) was driven by relative reductions in IRF and SNF payments. Average IRF payments decreased by $357 more for CJR episodes than for control group episodes from the baseline to the intervention period (p<0.01, Exhibit 11). This relative reduction in IRF payments equates to a 23.4% decrease in average IRF payments for CJR episodes from the baseline. Average SNF payments decreased by $508 more for CJR episodes than for control group episodes, or 9.3% from the CJR baseline (p<0.01, Exhibit 11).

Exhibit 11: The reduction in average episode payments was due to decreases in inpatient rehabilitation facility and skilled nursing facility payments, PY1-2

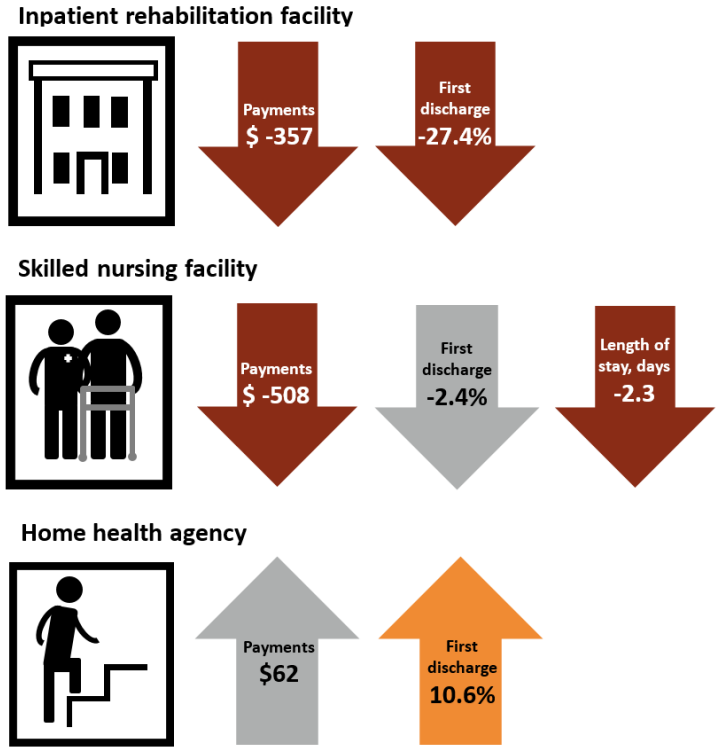


Source: Lewin analysis of Medicare claims and enrollment data for episodes initiated in 2012 through 2014 that ended between April 2012 and March 2015 (baseline) and episodes initiated during or after April 2016 that ended by December 2017 (intervention).

Notes: The estimates in this exhibit are the result of a difference-in-differences (DiD) model. DiD estimates that are significant at the 99%, 95%, or 90% significance level are indicated by dark, medium, and light orange shaded bars, respectively.

HHA = home health agency, IRF = inpatient rehabilitation facility, PY = performance year, SNF = skilled nursing facility.

Exhibit 12: The reasons for the decrease in post-acute care payments differ by setting



Source: Lewin analysis of Medicare claims and enrollment data for episodes initiated in 2012 through 2014 that ended between April 2012 and March 2015 (baseline) and episodes initiated during or after April 2016 that ended by December 2017 (intervention).

Notes: The estimated relative change in utilization is the result of a DiD model. DiD estimates that are significant at the 99%, 95%, or 90% significance level are indicated by dark, medium, and light orange shaded shapes, respectively.
The change in the proportion of patients first discharged to each PAC setting represents the percent change from the CJR baseline that is due to CJR. It is calculated by dividing the DiD estimate by the CJR baseline average. For IRF and SNF LOS, beneficiaries must have spent at least one day in the respective institutional PAC setting.

The relative decrease in IRF payments is the result of a relative reduction in the proportion of LEJR patients first discharged from the hospital to an IRF. The proportion of patients discharged to an IRF decreased more for CJR episodes than for control episodes, representing a 27.4% decrease from the baseline proportion ($p<0.01$, Exhibit 12).

During the baseline period, a greater proportion of CJR patients were discharged to an IRF than control group patients. The proportion of patients discharged to an IRF declined for both groups, although more so for CJR patients, so that during the intervention period, a smaller proportion of CJR than control patients were discharged to an IRF (Exhibit 13). We did not observe a relative change in the number of days that CJR patients spent in an IRF, among patients with an IRF stay (Appendix H). This is not unexpected because Medicare pays IRFs a per-discharge rate that does not vary by length of stay (LOS). Thus, the CJR model does not create an incentive to reduce IRF LOS.

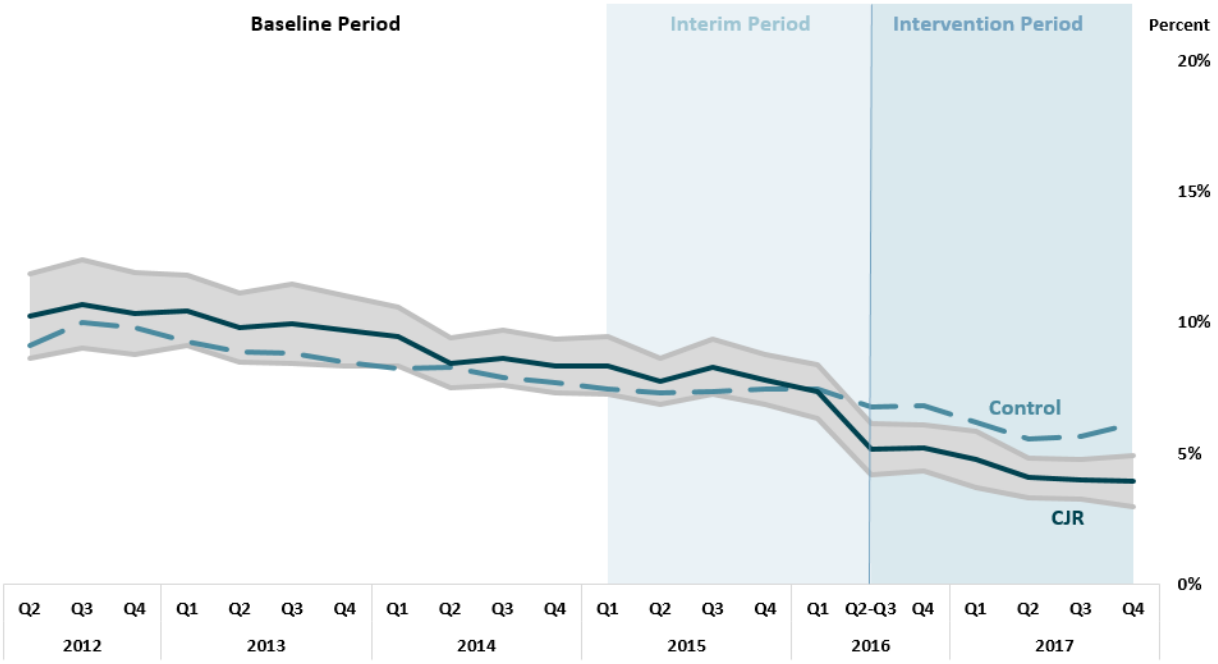
The relative decrease in SNF payments is the result of a relative decrease in the number of SNF days among patients with a SNF stay. The average number of SNF days decreased by 2.3 days more for CJR episodes than for control group episodes from the baseline to the intervention period ($p<0.01$, Exhibit 12). In contrast to how Medicare pays IRFs, Medicare pays SNFs a daily rate. Thus, hospitals have an incentive to influence SNFs to reduce the number of days of SNF care to lower episode payments.

There was no change in average HHA payments, even though the proportion of patients first discharged to HHA increased. This proportion increased more for CJR episodes than for control group episodes from the baseline to the intervention period, an increase of 10.6% from the CJR baseline proportion ($p<0.10$, Exhibits 12 and 14). Because some patients receive home health care

following discharge from institutional PAC, we also examined the proportion of patients who received care from an HHA at any time during their episode. There was no statistically significant change in the proportion of patients who had home health care at any time during the episode. This contributed to the lack of change in average HHA payments because HHA payments reflect HHA care received any time during the episode.

The relative decrease in the proportion of CJR patients discharged to IRF and the increase in the proportion discharged to HHA suggest that the CJR model resulted in shifts in care from more intensive to less intensive PAC settings. This result is consistent with the expectation that hospitals would respond to CJR by reducing the use of a costlier institutional PAC and replace it with less expensive HHA services to reduce payments during the episode of care.

Exhibit 13: The proportion of patients discharged to inpatient rehabilitation facilities declined more for CJR than for control episodes, PY1–2



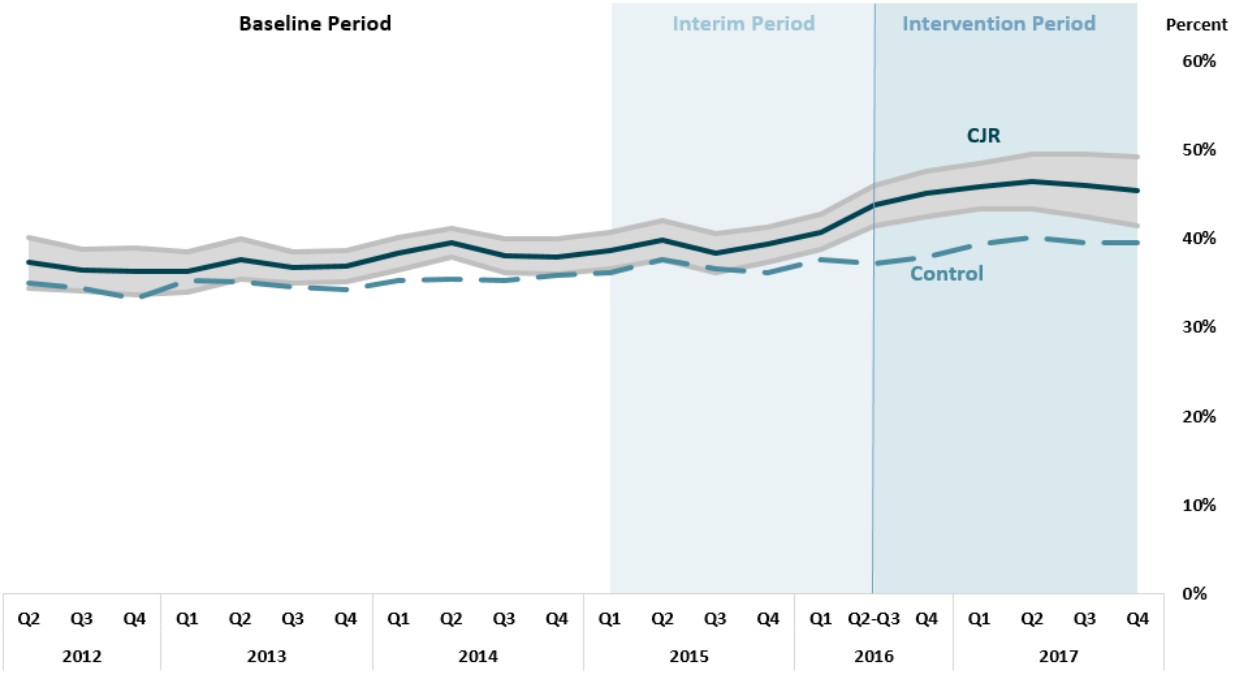
Source: Lewin analysis of Medicare claims and enrollment data for episodes initiated on or after January 2012 that ended by December 2017.

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

The gray shading represents the 95% confidence interval for the CJR estimate.

PY = performance year.

Exhibit 14: The proportion of patients first discharged to home health agencies increased more for CJR episodes than for control episodes, PY1-2



Source: Lewin analysis of Medicare claims and enrollment data for episodes initiated on or after January 2012 that ended by December 2017.

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

The gray shading represents the 95% confidence interval for the CJR estimate.

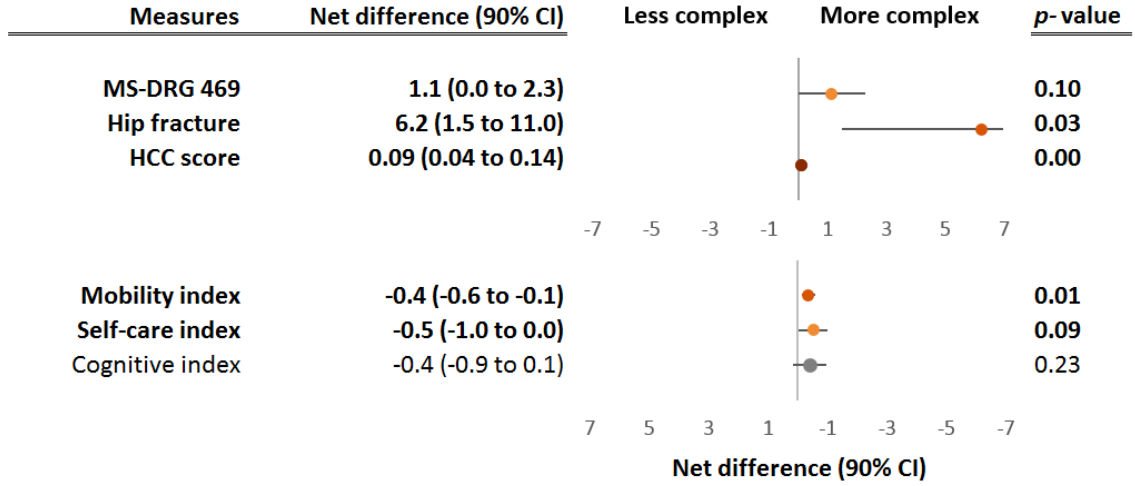
PY = performance year.

b. Were changes in patient mix by setting consistent with changes in utilization?

With shifts from more intensive to less intensive PAC settings, we would expect patient complexity to increase in the IRF, SNF, and HHA settings. There is evidence this occurred in IRFs, some evidence to suggest this may have occurred in SNFs, but no evidence that it did in HHAs.

The complexity of the average CJR patient discharged to an IRF was greater during the intervention period than the baseline period, relative to the control group, based on five out of six measures of patient complexity. There were relative decreases in the average mobility index ($p<0.05$) and self-care index ($p<0.10$) for CJR patients. There were also relative increases in the proportion of CJR patients with a fracture ($p<0.05$) and the proportion discharged under the more complex MS-DRG 469 ($p<0.10$), as well as a relative increase in CJR patients' average HCC score ($p<0.01$) (Exhibit 15).

Exhibit 15: Patient complexity increased for CJR patients first discharged to an inpatient rehabilitation facility, PY1-2



Source: Lewin analysis of Medicare claims, enrollment, and Inpatient Rehabilitation Facility Patient Assessment Instrument data for episodes initiated in 2012 through 2014 that ended between April 2012 and March 2015 (baseline) and episodes initiated during or after April 2016 that ended by September 2017 (intervention).

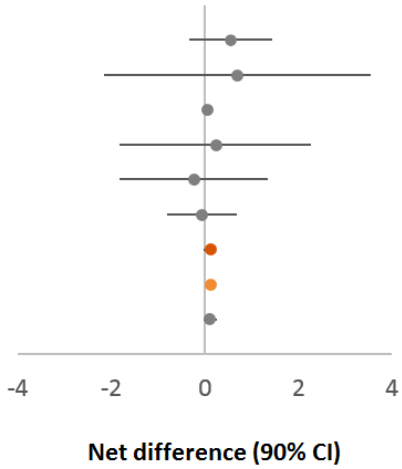
Notes: Net differences that are significant at the 99%, 95%, or 90% significance level are indicated by dark, medium, and light orange shaded points, respectively.

CI = confidence interval, HCC = hierarchical condition category, MS-DRG = Medicare Severity-Diagnosis Related Group, PY = performance year.

There was also some evidence of an increase in the complexity of CJR patients who were discharged to a SNF, relative to control group patients. Changes in two of nine complexity measures indicated a statistically significant relative decrease in CJR patients' functional status at SNF admission. The relative increases in CJR patients' average early-loss activities of daily living (ADLs) scores ($p<0.05$) and motion scores ($p<0.10$) suggest an increase in patients with greater needs were discharged to a SNF relative to the control group (Exhibit 16).

Exhibit 16: For patients first discharged to skilled nursing facilities, some evidence of an increase in patient complexity, PY1–2

Measures	Net difference (90% CI)	*p*-value
MS-DRG 469	0.5 (-0.3 to 1.4)	0.31
Hip fracture	0.7 (-2.2 to 3.5)	0.69
HCC score	0.04 (-0.01 to 0.09)	0.16
Bathing poor	0.2 (-1.8 to 2.3)	0.85
Cognition not intact	-0.2 (-1.8 to 1.4)	0.81
Severe cognitive impairment	-0.1 (-0.8 to 0.7)	0.87
Early-loss ADL score	**0.1 (0.0 to 0.2)**	**0.02**
Motion score (i.e., mid-loss ADL score)	**0.1 (0.0 to 0.2)**	**0.07**
Late-loss ADL score	0.1 (0.0 to 0.3)	0.17



Source: Lewin analysis of Medicare claims, enrollment and Minimum Data Set (MDS) data for episodes initiated in 2012 through 2014 that ended between April 2012 and March 2015 (baseline) and episodes initiated during or after April 2016 that ended by September 2017 (intervention).

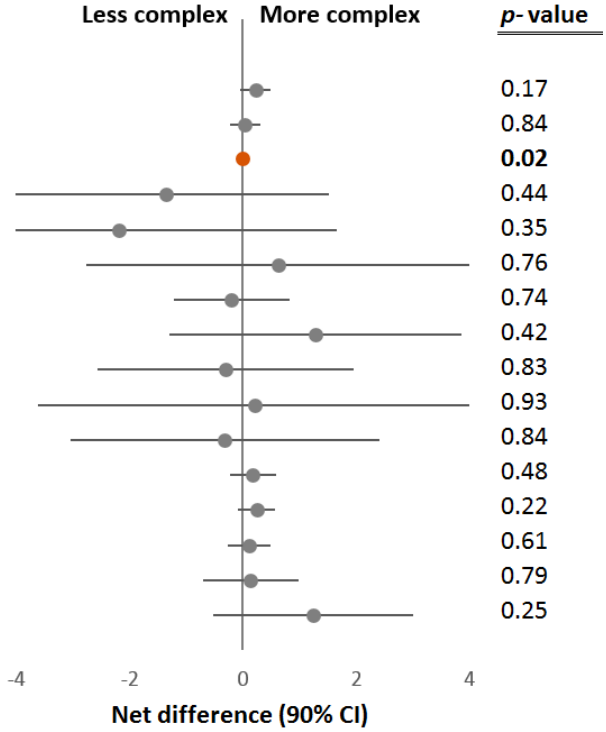
Notes: Net differences that are significant at the 99%, 95%, or 90% significance level are indicated by dark, medium, and light orange shaded points, respectively.

ADL = activities of daily living, CI = confidence interval, HCC = hierarchical condition category, MS-DRG = Medicare Severity-Diagnosis Related Group, PY = performance year.

There was no consistent evidence that CJR patients first discharged to HHA were relatively more complex than control patients in the intervention period. Only one of 16 patient characteristics suggested a relative increase in complexity: there was a small but statistically significant increase of 0.02 index points in the average HCC score for CJR patients relative to control patients ($p<0.05$, Exhibit 17).

Exhibit 17: No consistent evidence of a relative increase in patient severity for CJR patients first discharged to home health, PY1–2

Measures	Net difference (90% CI)	*p*-value
MS-DRG 469	0.2 (0.0 to 0.5)	0.17
Hip fracture	0.0 (-0.2 to 0.3)	0.84
HCC score	**0.02 (0.01 to 0.03)**	**0.02**
Toilet transferring	-1.4 (-4.2 to 1.5)	0.44
Transferring	-2.2 (-6.0 to 1.6)	0.35
Ambulation / locomotion	0.6 (-2.8 to 4.0)	0.76
Lower body dressing	-0.2 (-1.2 to 0.8)	0.74
Upper body dressing	1.3 (-1.3 to 3.9)	0.42
Bathing	-0.3 (-2.6 to 2.0)	0.83
Toileting hygiene	0.2 (-3.6 to 4.0)	0.93
Grooming	-0.3 (-3.1 to 2.4)	0.84
Cognitive functioning	0.2 (-0.2 to 0.6)	0.48
Confusion	0.2 (-0.1 to 0.6)	0.22
Memory deficit	0.1 (-0.3 to 0.5)	0.61
Impaired decision-making	0.1 (-0.7 to 1.0)	0.79
Overall status	1.2 (-0.5 to 3.0)	0.25



Source: Lewin analysis of Medicare claims, enrollment and Outcome and Assessment Information Set (OASIS) data for episodes initiated in 2012 through 2014 that ended between April 2012 and March 2015 (baseline) and episodes initiated during or after April 2016 that ended by September 2017 (intervention).

Notes: Net differences that are significant at the 99%, 95%, or 90% significance level are indicated by dark, medium, and light orange shaded points, respectively.

CI = confidence interval, HCC = hierarchical condition category, MS-DRG = Medicare Severity-Diagnosis Related Group, PY = performance year.

c. *Conclusion*

CJR participant hospitals decreased LEJR episode payments by reducing the use of the more intensive institutional PAC. The proportion of LEJR patients who received IRF care decreased and for those who received SNF care, the number of days in the SNF went down. Both of these changes contributed to lower episode payments. The proportion of patients who were discharged from the hospital to an HHA went up, although average HHA payments per episode did not change. The average complexity of LEJR patients who received PAC in an IRF increased, indicating that the less intensive patients were less likely to receive IRF care under the CJR model.

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

The CJR model was designed to reward hospitals that delivered high quality of care for Medicare beneficiaries undergoing LEJR. To encourage participant hospitals to focus on improving quality, a lower discount is applied to the target price for participant hospitals with a higher quality score (which results in a higher quality-adjusted target price).

1. Key findings

- Quality of care, as measured by the unplanned readmission rate, emergency department use, and mortality, was maintained under the CJR model.

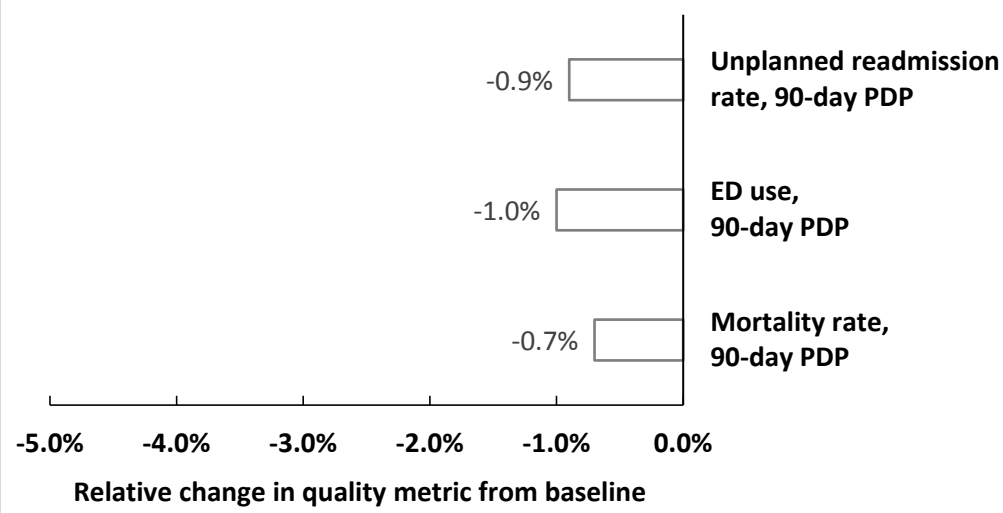
2. Methods

This analysis used the DiD modeling approach described in Section II.A.2 to estimate relative change in outcomes. For more detailed information about the methodology, see Appendix E.

3. Results

During the first two performance years, the CJR model had no statistically significant impact on the unplanned readmission rate, emergency department (ED) use, or the mortality rate during the 90-day post-discharge period. Changes in these claims-based quality metrics from baseline were not statistically different for CJR episodes and control group episodes (Exhibit 18).

Exhibit 18: Quality of care was maintained under the CJR model, PY1-2



Source: Lewin analysis of Medicare claims and enrollment data for episodes initiated in 2012 through 2014 that ended between April 2012 and March 2015 (baseline) and episodes initiated during or after April 2016 that ended by December 2017 (intervention).

Notes: The estimates in this exhibit are the result of a difference-in-differences (DiD) model. DiD estimates that are significant at the 99%, 95%, or 90% significance level are indicated by dark, medium, and light orange shaded bars, respectively.

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

ED = emergency department, PDP = post-discharge period, PY = performance year.

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

As discussed in the impact of the CJR model on quality of care section (Section II.E), the CJR model is intended to encourage participant hospitals to improve quality of care by coordinating care with the physicians, PAC providers, and other providers and clinicians involved in the episode. At the same time, there is a concern that the incentive to lower episode payments could result in changes in care that lead to poorer long term outcomes. Thus, the examination of functional status and pain at later stages of the recovery period are of particular importance for understanding the impact of the CJR model.

Functional status, pain, and care experience are indicative of quality of care, but cannot be measured through Medicare claims data. As such, we rely on a patient survey, which reflects patient recovery after the end of the episode. We also examined patient assessments conducted during PAC stays, which provide interim measures of functional status and pain.

1. Key findings

- After the end of the episode, CJR and control survey respondents had similar self-reported changes in functional status and pain.
- For patients discharged from the hospital to an IRF, SNF, or HHA, a smaller proportion of CJR patients generally improved their functional status than control patients while in the PAC setting.
- There were no differences between CJR and control survey respondents in overall satisfaction with recovery, care management, or care transitions, but CJR survey respondents were slightly more reliant on caregivers after returning home.

2. Methods

a. Patient survey

We surveyed beneficiaries after the end of their LEJR episode to determine if CJR patients differed from control patients on several patient-reported outcomes. Measures included change in functional status and pain from the week before their hospitalization through the time of the survey (which was typically after the end of the CJR episode),[19] overall satisfaction with recovery, satisfaction with care management, care experiences, and caregiver help needed after returning home. We estimated risk-adjusted differences between CJR and control respondents that accounted for beneficiary, hospital, and MSA attributes. See Appendix E for more detail on the methods and Appendix I for the patient survey questions.

[19] We mailed surveys to patients between 60 and 120 days after their LEJR discharge. The median time at which surveys were returned was 43 days after the conclusion of the patient's 90-day post-discharge period.

We used a stratified random sampling design to select beneficiaries for the survey (Appendix E for more detail). Beneficiaries in the control groups were selected such that they were similar to the CJR sample according to hospital characteristics, such as historical LEJR volume and historical episode payments, and beneficiary characteristics (sex and age). Data were collected in two waves that covered episodes with hospital discharges in March or April 2017, and in September or October 2017. The overall response rate was approximately 71.6% in the CJR group and 72.2% in the control group. Survey results were based on 5,374 completed survey responses from CJR patients, which account for 19.2% of all CJR episodes in the four months covered by the survey. Although each wave was large enough to detect meaningful differences between CJR and control respondents, we pooled responses from both waves for this analysis to better capture the average across the second performance year.

b. Assessment-based measures

We conducted a DiD analysis to estimate the differential change in several functional status and pain measures for patients discharged to IRF, SNF, or HHA. These measures are captured from required, comprehensive assessments completed at the start and end of each PAC stay. The assessment instrument varies by setting; measures are based on the Inpatient Rehabilitation Facility Patient Assessment Instrument (IRF-PAI) (for IRF patients), the Minimum Data Set (MDS) (for SNF patients), and the Outcome and Assessment Information Set (OASIS) (for HHA patients). The measures vary across the settings, as does the timing between the admission and discharge assessments. While this precludes a direct comparison of the measures across settings, the measures provide valuable information about changes in functional status and pain within each setting. We report on one functional status measure for patients initially discharged to an IRF, two functional status measures and one pain measure for those initially discharged to a SNF, and two functional status measures and one pain measure for those initially discharged to an HHA. This analysis relies on the same baseline period as the claims-based analyses, but the intervention period is one quarter shorter (from April 2016 to September 2017) because of the longer time needed for assessment data to become available.[20] The results are risk-adjusted to control for functional status at the initiation of PAC; health care service use before the anchor hospitalization; and beneficiary, market, and hospital characteristics. Risk adjusting the estimates is particularly important because the CJR model has affected the initial discharge setting for LEJR patients as well as the length of stay within these settings.[21] Details about the DiD estimator and risk adjustment models are discussed in Appendix E.

3. Results

a. Functional status and pain

This section presents the results from the patient survey and the PAC assessments on patient functional status and pain. The survey data capture changes in functional status later in the course

[20] While this intervention period is three months shorter than the intervention period for other analyses in this report, this does not appear to affect our conclusions because the results have been stable over time.

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

of recovery, whereas the PAC assessments represent gains made during the PAC stay, which typically ends earlier in the recovery period.

Patient survey results

The patient survey provides information across patients regardless of PAC setting, and includes those who were discharged home without HHA care. Additionally, the survey measures functional status and pain later in the recovery period than the PAC assessments.

Survey respondents were asked eight questions regarding their functional status and pain. Overall, respondents in the CJR and control groups reported improvement in all eight measures of functional status and pain from before to after their surgery (Exhibit 19). Differences in the amount of change between the CJR and control groups were small and not statistically significant.[22] These results indicate that patients experienced similar improvements in function and level of pain, regardless of whether they received LEJR in a CJR participant hospital or a control group hospital.

Exhibit 19: CJR and control respondents experienced similar levels of improvement in functional status and pain

Survey measure	Change in self-reported measure from before the hospitalization to after the episode *Higher value represents a more favorable change*		Difference between CJR and control groups
	CJR	Control group	
Ability to walk by yourself without resting[a]	0.82	0.79	0.03
Difficulty walking up or down 12 stairs[b]	0.83	0.82	0.00
Difficulty rising from sitting[a]	1.28	1.26	0.01
Difficulty standing[a]	1.23	1.22	0.01
Use of a mobility aid[c]	0.19	0.22	-0.02
Difficulty getting on/off the toilet[a]	1.43	1.41	0.02
Frequency that pain interferes with normal activities[a]	2.00	2.05	-0.04
Medication use for pain in the joint you had replaced[b]	0.60	0.63	-0.03

Source: Lewin analysis of patient survey data for episodes with discharge in March, April, September, and October 2017.

Notes: The change in a given measure refers to the difference between a respondent's self-reported status at the time of the survey and the respondent's recalled pre-hospital status. Estimated changes, and the difference between changes in the CJR and control group, are reported in "level" terms.

The estimates in this exhibit are the results of a cross-sectional regression model, weighted for sampling and nonresponse. Estimates that are significant at the 99%, 95%, or 90% significance level are indicated by dark, medium, and light orange shaded bars, respectively. There were no significant differences between CJR and control respondents in the measures reported in this exhibit.

[a] Indicates the question has 5 possible responses (i.e., "levels"), and the change could range from -4 to 4.

[b] Indicates the question has 4 possible responses (i.e., "levels"), and the change could range from -3 to 3.

[c] Indicates the question has 3 possible responses (i.e., "levels"), and the change could range from -2 to 2.

[22] If CJR and control respondents had substantially different pre-surgery functional status, then changes in functional status from before the hospitalization to the time of the survey would not be measured on the same scale in the two groups, which would yield biased results. We did not find any evidence that CJR and control respondents had different pre-hospital functional status (Appendix K).

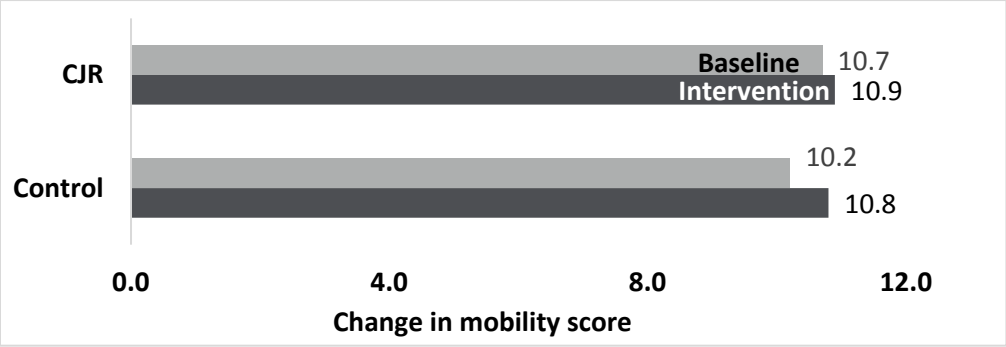
Assessment-based results

PAC assessment data measure changes in functional status and pain for patients discharged from the hospital to an IRF, SNF, or HHA. The data capture changes made between the beginning and end of a patient's first PAC stay.

For patients first discharged to IRF. IRF care is typically more intensive than care in other PAC settings and Medicare payments are typically higher. Less than 10% of CJR patients were discharged to an IRF during the baseline period. IRF stays averaged 11 days.

For CJR patients first discharged to an IRF, the average change in mobility score during their IRF stay improved by 0.3 points from baseline to intervention (from 10.7 to 10.9); for the control group the average change in mobility score improved by 0.6 points (from 10.2 to 10.8) (Exhibit 20). As a result, the average change in mobility score for CJR patients decreased by 0.3 points relative to the control group ($p<0.10$).[23] Although the average change in mobility score for CJR patients did not increase as much as it did for control patients from the baseline to the intervention period, the CJR and control patients discharged to an IRF had comparable improvements in mobility score during the intervention period.

Exhibit 20: The change in mobility score from baseline to intervention was less for CJR than control patients first discharged to an IRF, although the change in mobility score for CJR and control patients was comparable during the intervention period, PY1–2



Source: Lewin analysis of Inpatient Rehabilitation Facility Patient Assessment Instrument (IRF-PAI) data for episodes initiated in 2012 through 2014 that ended between April 2012 and March 2015 (baseline) and episodes initiated during or after April 2016 that ended by September 2017 (intervention).

Notes: The change in mobility score ranges from -23 to 24. A positive change in mobility score indicates improvement.

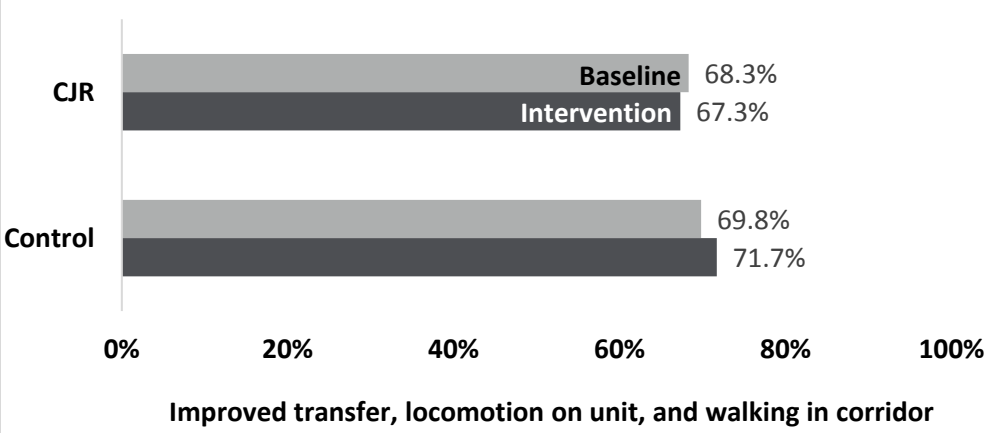
PY = performance year. IRF = inpatient rehabilitation facility.

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

For patients first discharged to SNF. SNF care is typically less intense than IRF care and Medicare payments are generally lower. Almost 40% of CJR LEJR patients were discharged to SNFs during the baseline period and their stays averaged about 25 days. Under the CJR model, SNF average length of stay decreased 2.3 days (see below for a sensitivity analysis of the impact of reduced time in the SNF on functional status improvements).

The majority of CJR patients first discharged to a SNF improved in mobility, which was measured as transfer, locomotion on unit, and walking in corridor, during their SNF stay. The proportion of CJR patients with improvement in mobility decreased by 1.0 percentage point from baseline to intervention (from 68.3% to 67.3%). For the control group, the proportion of patients with improvement in mobility increased by 1.9 percentage points (from 69.8% to 71.7%) (Exhibit 21). As a result, the proportion of CJR patients with improvement in mobility decreased by 2.9 percentage points relative to the control patients ($p<0.05$, Appendix H).

Exhibit 21: The proportion of patients with improved mobility from baseline to intervention decreased for CJR patients discharged to SNF, compared with an increase for control patients, PY1-2.

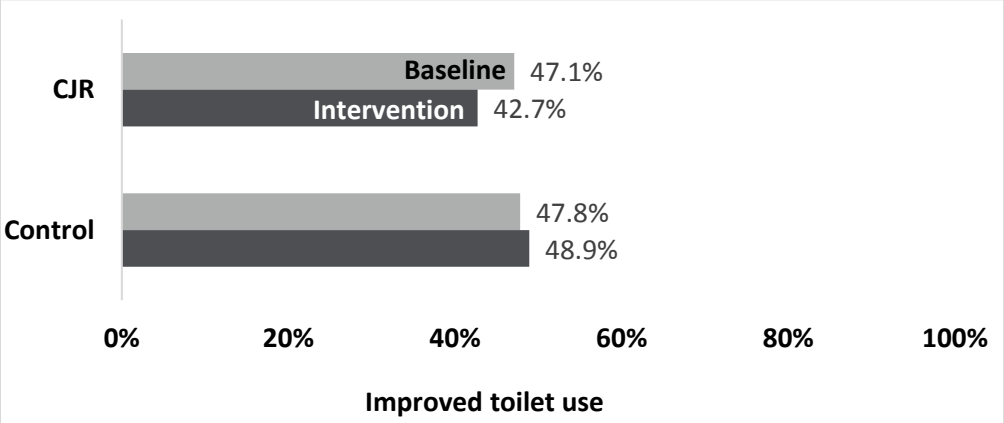


Source: Lewin analysis of Minimum Data Set (MDS) data for episodes initiated in April 2012 and March 2015 (baseline) and episodes initiated during or after April 2016 that ended by September 2017 (intervention).

Notes: PY = performance year, SNF = skilled nursing facility.

Less than half of CJR patients first discharged to a SNF improved in toilet use during their SNF stay, and the proportion with improved toilet use decreased by 4.5 percentage points from baseline to intervention (from 47.1% to 42.7%); for the control group it increased by 1.2 percentage points (from 47.8% to 48.9%) (Exhibit 22). As a result, the proportion of CJR patients first discharged to a SNF who improved in toilet use decreased by 5.6 percentage points relative to the control patients, or 11.9% from the CJR baseline ($p<0.05$, Appendix H).

Exhibit 22: The proportion of patients with improved toilet use from baseline to intervention decreased for CJR patients first discharged to a SNF, compared with an increase for control patients, PY1-2.



Source: Lewin analysis of Minimum Data Set (MDS) data for episodes initiated in April 2012 and March 2015 (baseline) and episodes initiated during or after April 2016 that ended by September 2017 (intervention).

Notes: These results need to be interpreted with caution because they do not pass the parallel trend test.

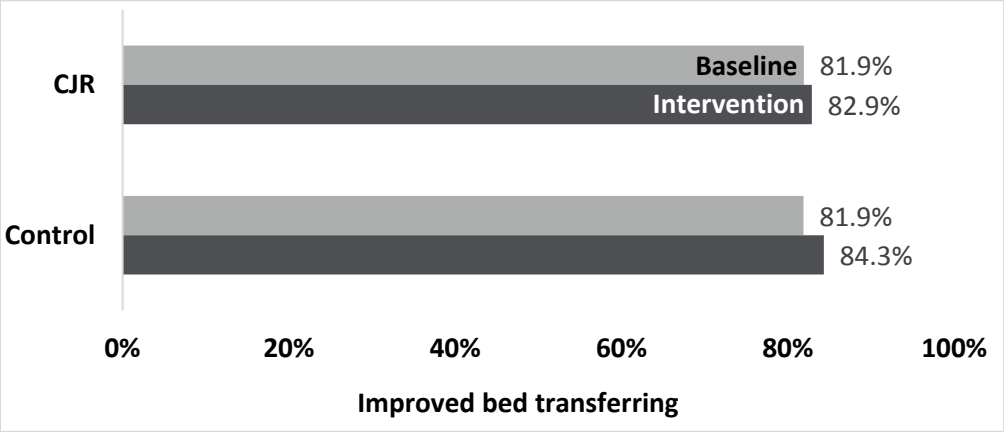
PY = performance year, SNF = skilled nursing facility.

There was no statistical difference in the proportion of patients without moderate or severe pain between CJR and control patients first discharged to a SNF from the baseline to the intervention (Appendix H).

For patients first discharged to HHA. HHA care is typically the least intensive PAC setting. Almost 40% of CJR patients were first discharged to an HHA. The average time between patient assessments taken at the beginning and end of care was about 23 days during the baseline so, as with the IRF and SNF measures, the changes examined represent early functional gains.

The majority of CJR patients first discharged home with HHA care improved in bed transferring. The proportion of CJR patients with improvement in bed transferring increased by 1.0 percentage point from baseline to intervention (from 81.9% to 82.9%). For the control group, the proportion increased by 2.4 percentage points (from 81.9% to 84.3) (Exhibit 23). As a result, the proportion of CJR patients with improvement in bed transferring decreased relative to control patients by 1.4 percentage points, or 1.7% from the CJR baseline ($p<0.05$, Appendix H).

Exhibit 23: The proportion of patients with improved bed transferring from baseline to intervention increased for CJR patients first discharged to HHA, but increased more for control patients, PY1-2.

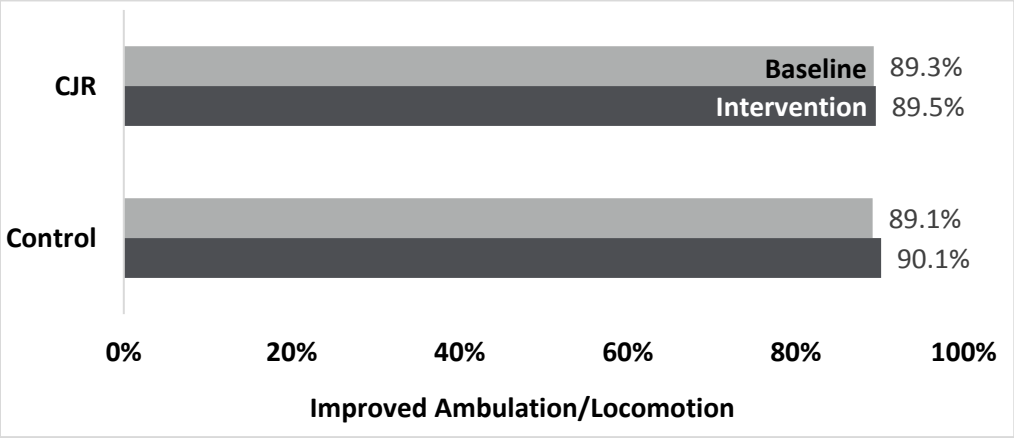


Source: Lewin analysis of Outcome and Assessment Information Set (OASIS) data for episodes initiated in April 2012 and March 2015 (baseline) and episodes initiated during or after April 2016 that ended by September 2017 (intervention).

Notes: HHA= home health agency, PY = performance year.

The majority of CJR patients first discharged home with HHA care improved in ambulation/ locomotion, and the proportion was nearly the same in the baseline and intervention periods (89.3% and 89.5%). For the control group, the proportion increased by 1.0 percentage point (from 89.1% to 90.1%) (Exhibit 24). As a result, the proportion of CJR patients who improved in motion decreased relative to control patients by 0.8 percentage points ($p<0.10$, Appendix H). Nearly all of the patients improved for both groups and the relative differences are small.

Exhibit 24: The proportion of patients with improved ambulation/locomotion from baseline to intervention increased for CJR patients first discharged to HHA, but increased more for control group patients, PY1-2.



Source: Lewin analysis of Outcome and Assessment Information Set (OASIS) data for episodes initiated in April 2012 and March 2015 (baseline) and episodes initiated during or after April 2016 that ended by September 2017 (intervention).

Notes: HHA= home health agency, PY = performance year.

There was no statistical difference in the proportion of patients with improvement in pain interfering with activity between CJR and control patients first discharged to an HHA from the baseline to the intervention (Appendix H).

Sensitivity findings

We conducted additional analyses to better understand the changes in functional status of CJR patients who were discharged to PAC relative to the control group. The CJR model resulted in relative reductions in the length of SNF stays (-2.3 days, $p<0.01$) and the number of HHA visits (-0.8 visits, $p<0.01$). This indicates that CJR patients in these settings likely had less time for improvements in functional status and pain while receiving PAC. The CJR model did not result in changes in IRF days (Section II.D.3). To understand whether less time in the SNF or HHA settings contributed to the relative declines in the proportion of patients with improvement, we measured the average time between assessments for patients discharged to a SNF or HHA, and we re-

calculated the pain and functional status estimates controlling for the number of days between assessments.[24]

Consistent with the relative decrease in the number of SNF days, the number of days between patient assessments decreased by 1.9 days more for CJR patients ($p<0.01$, Appendix J). After controlling for the number of days between patient assessments in the risk adjustment models, the estimated impact of the CJR model on mobility decreased from a 4.2% decrease to a 1.6% decrease and was no longer statistically significant (Appendix J). The estimated impact of the CJR model on toilet use decreased from an 11.9% decrease to a 9.0% decrease from the CJR baseline, but remained statistically significant ($p<0.05$, Appendix J). Thus, for CJR patients first discharged to SNF, the decrease in the number of days between assessments explains little of the relative decrease in the proportion of patients with improvement in toilet use.

For patients first discharged home with HHA, the number of days between patient assessments decreased by 2.0 days more for CJR patients ($p<0.01$, Appendix J), consistent with the relative decrease in HHA visits. After controlling for the number of days between patient assessments, the estimated impact of the CJR model was reduced by less than 0.2 percentage points and the estimates were less statistically significant (Appendix J). Thus, for CJR patients first discharged home with HHA, the decrease in the number of days between assessments explains little of the relative decrease in the proportion of patients with improvement in functional status.

In another sensitivity analysis, we found that the majority of LEJR patients with no improvement in their functional status during their PAC stay maintained their functional status, that is, they did not decline in their status. (Appendix J).

b. Satisfaction, care transitions, and caregiver help

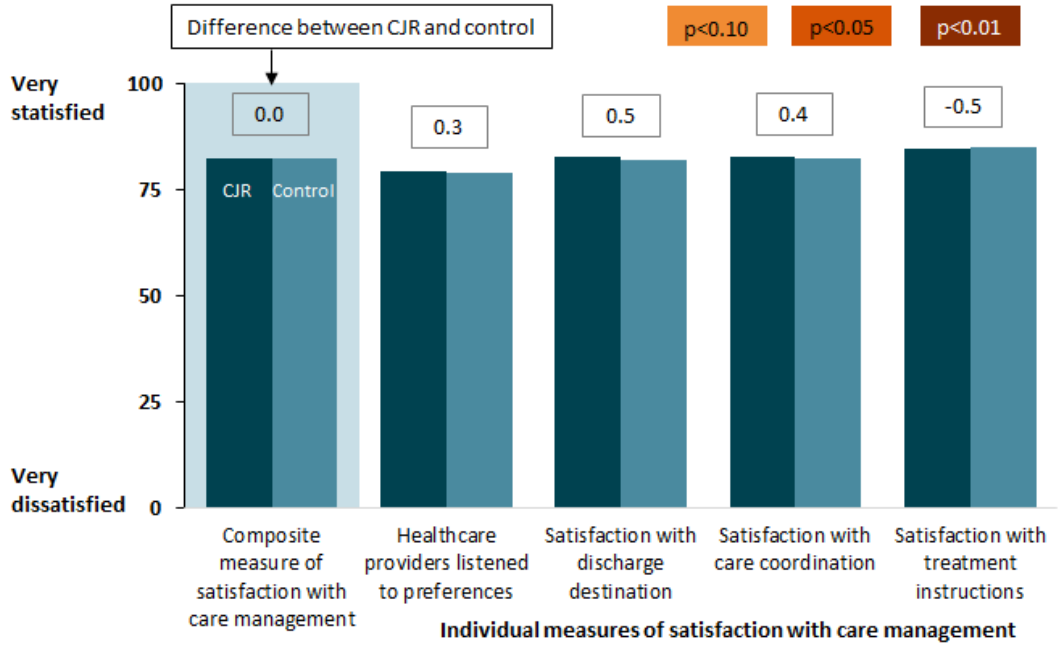
The patient survey also captured overall satisfaction with recovery, satisfaction with care management, care experiences, and caregiver help needed after returning home.

[24] Because the CJR model may impact both the length of SNF stays, number of HHA visits, and ADL outcomes, the number of days between assessments is not included as a causal risk factor in the risk adjustment models for the main analysis. However, including the variable in the risk adjustment models helps us determine whether the relative change in the proportion of CJR patients that improved in functional status can be explained by its association with the relative change in number of days between assessments among CJR patients.

Satisfaction

On average, both CJR and control respondents were satisfied with their overall physical recovery (Appendix K) and with the way their care was managed (Exhibit 25). There were no statistically significant differences between CJR and control respondents on any measure of satisfaction related to physical recovery (Appendix K) or care management (Exhibit 25).

Exhibit 25: CJR and control respondents were similarly satisfied with their care management



Source: Lewin analysis of patient survey data for episodes with discharge in March, April, September, and October 2017.

Notes: Differences that are significant at the 99%, 95%, or 90% significance level are indicated by dark, medium, and light orange shading, respectively.

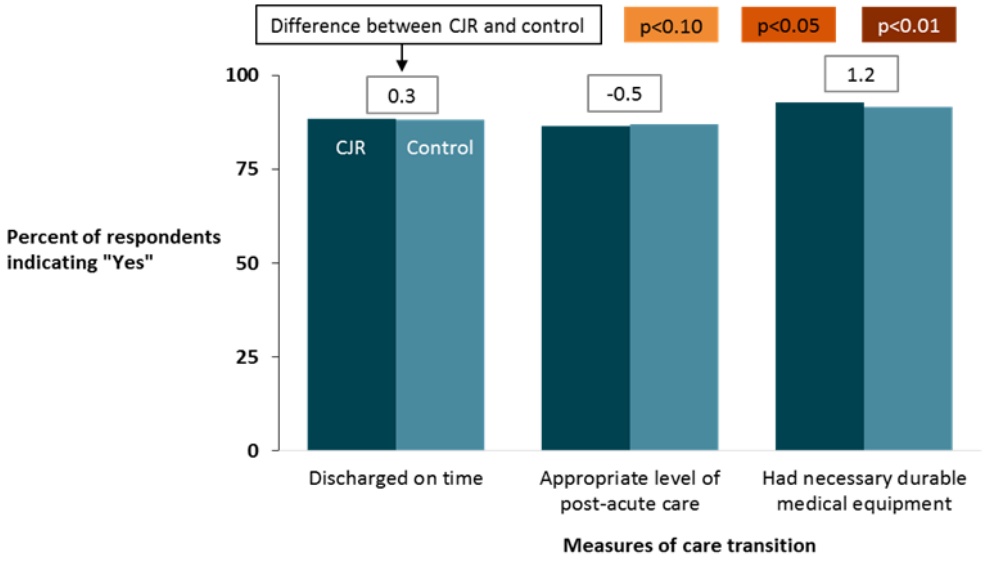
All outcomes are scaled from 0 to 100 points, where 0 = very dissatisfied, 25 = somewhat dissatisfied, 50 = neutral, 75 = somewhat satisfied, and 100 = very satisfied.

The composite summarizes all four measures of satisfaction with care management. Differences between CJR and control outcomes are reported in point terms.

Care Transitions

Overall, both the CJR and control group respondents indicated positive experiences with care transitions. Over 85% indicated that they were discharged on time, received appropriate post-discharge care, and had access to durable medical equipment (Exhibit 26). There were no statistically significant differences between CJR and control respondents.

Exhibit 26: CJR and control respondents reported similarly positive care transitions



Source: Lewin analysis of patient survey data for episodes with discharge in March, April, September, and October 2017.

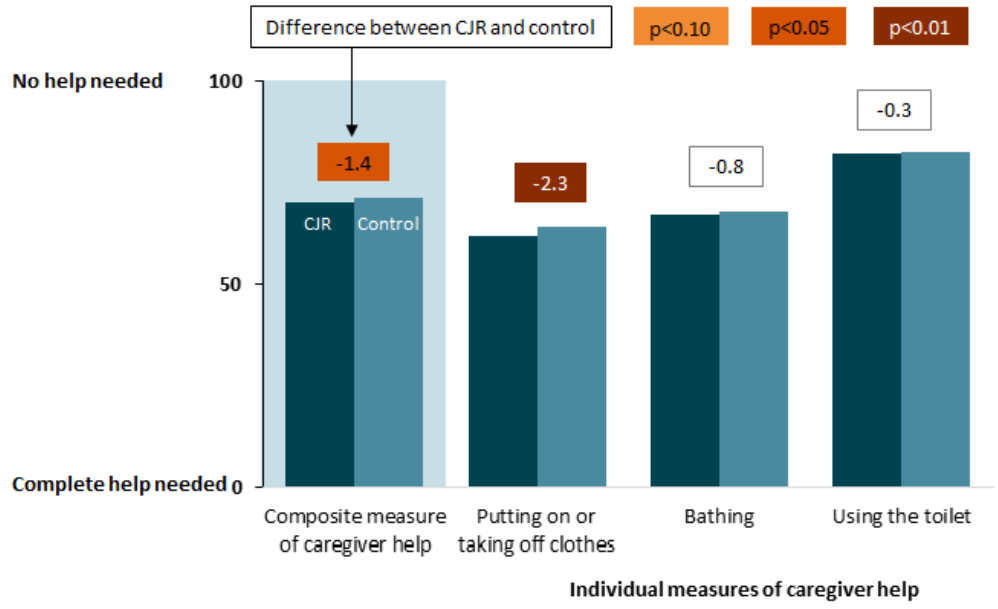
Notes: Differences that are significant at the 99%, 95%, or 90% significance level are indicated by dark, medium, and light orange shading, respectively.

All outcomes are in percentage terms, ranging from 0 to 100. Differences between CJR and control outcomes are reported in percentage point terms.

Caregiver help

Approximately 95% of CJR and control survey respondents received at least some help from a caregiver after returning home, and differences between the groups for this measure were not statistically significant (Appendix K). Among those receiving help, CJR respondents reported needing more help than control respondents based on a composite measure of three ADLs ($p<0.05$, Exhibit 27). Differences in the composite score were primarily driven by a measure of needing caregiver help to put on and take off clothes: CJR respondents scored 2.3 points lower on this measure ($p<0.01$) on a 100-point scale: a small difference. Our clinical review panel indicated that the magnitude of the difference was neither concerning nor clinically relevant. Differences between CJR and control respondents in the amount of help needed bathing or using the toilet were not statistically significant.

Exhibit 27: CJR respondents required slightly more help



Source: Lewin analysis of patient survey data for episodes with discharge in March, April, September, and October 2017.

Notes: Differences that are significant at the 99%, 95%, or 90% significance level are indicated by dark, medium, and light orange shading, respectively.

Respondents were only asked about the amount of help needed with a given activity of daily living if they indicated that they received caregiver help.

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

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

c. Conclusion

The self-reported outcomes from the patient survey indicate that CJR and control patients had similar levels of improvement in functional status from the week prior to the surgery until after the

end of the episode. This indicates that longer-term functional status and pain outcomes were not adversely affected under the CJR model, despite changes in care delivery.

We also examined shorter-term changes in functional status and pain for patients discharged to an IRF, SNF, or HHA. These assessment-based outcomes generally indicated a decrease in the proportion of CJR patients who improved their functional status from the baseline to the intervention compared to the control group. For patients discharged to an IRF or HHA, the estimates for CJR and for the control group are similar in the intervention period, and both groups showed improvement from baseline to intervention. For patients discharged to a SNF, the proportion of CJR patients whose functional status improved decreased. By contrast, there was an increase in the proportion of control group patients whose functional status improved.

CJR patients discharged from the hospital to a SNF or HHA had less time to improve their functional status while receiving PAC because they spent relatively less time in the SNF or HHA setting. The decrease in the number of days between assessments, however, explained little of the relative decrease in the proportion of patients with improvement in functional status.

These PAC assessment results could reflect increased patient complexity in each PAC setting as the CJR model shifts PAC services from more intense PAC settings to less intense PAC settings. Although we risk adjusted our estimates to control for differences in patient mix, there could be changes in unobserved patient characteristics that we are not capturing in our models.

The self-reported outcomes from the patient survey show that there were no differences between CJR and control survey respondents in overall satisfaction with recovery, satisfaction with care management, or care transitions. However, CJR survey respondents reported a slightly higher reliance on caregivers than control respondents. The clinical experts who reviewed these findings concluded that the statistically significant differences were not clinically meaningful.

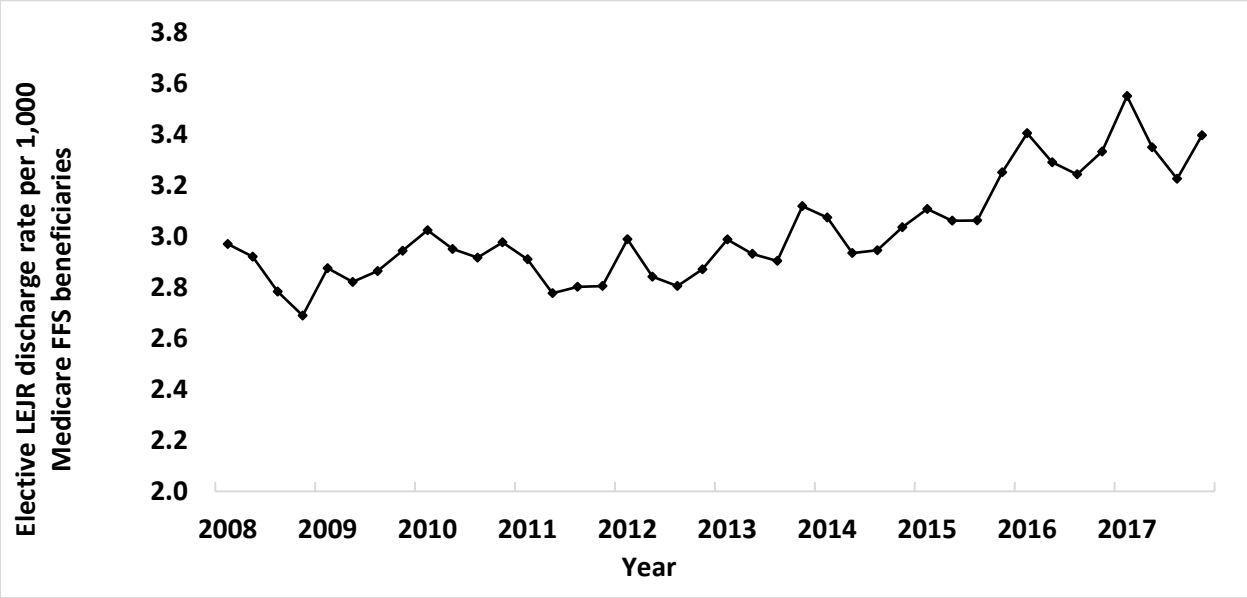
G. Did the CJR model result in any unintended consequences?

While the CJR model encourages participant hospitals to reduce episode payments and improve quality by coordinating care with the physicians, PAC providers, and other providers and clinicians involved in the episode, participant hospitals could achieve these goals through other means. One unintended consequence of the model would be if, rather than redesign care delivery, participant hospitals pursued alternative ways to lower episode payments or improve quality, thus increasing the likelihood of earning reconciliation payments. For instance, they could attempt to lower episode payments or improve quality scores by increasing episode volume, changing patient mix, or delaying services until after the episode.

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

The mandatory CJR model targets LEJR surgery in part because of its prevalence in the Medicare population, with more than 400,000 procedures in 2014 and growth projected to continue.[25] The volume of these procedures has been trending upward since the 1990s, with rates of total hip and knee replacements approximately doubling among those 45 and older between 2000 and 2010 (Exhibit 28).[26, 27]

Exhibit 28: Increasing national trend in the elective LEJR discharge rate since 2007



Source: Lewin analysis of Medicare claims and enrollment data for discharges from October 2007 through December 2017.

Notes: LEJR = lower extremity joint replacement.

In the context of overall growth in LEJR procedures, there are concerns that the CJR model itself could boost LEJR volume beyond current trends by making the procedure more financially rewarding to participant hospitals and their partners. Participant hospitals may be able to reduce average episode payments by providing LEJR to beneficiaries who otherwise would have foregone the procedure, because these are likely to be beneficiaries with fewer health needs and, therefore, less costly episodes. Medicare savings would be offset by the payments for these additional episodes.

---

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

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

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

CJR participant hospitals may also increase their volume of LEJR episodes if they can shift episodes from other hospitals through enhanced marketing, higher quality, or new gainsharing agreements with referring physicians. Shifts in volume across providers, however, would likely not have much effect on Medicare savings.[28]

a. Key findings

- The CJR model did not have a statistically significant impact on the volume of elective LEJR discharges in mandatory CJR markets.

b. Methods

We analyzed the impact of the CJR model on the volume of elective LEJR discharges in a market by estimating the relationship between CJR "dose" and the change in the elective LEJR discharge rate (discharges per 1,000 Medicare FFS beneficiaries) in MSAs. The Bundled Payments for Care Improvement initiative had an LEJR episode that was structured similarly to CJR episodes, and therefore had similar incentives for boosting volume of elective LEJR surgeries. Some participant hospitals were taking part in the Bundled Payments for Care Improvement initiative before and during the CJR model, so we controlled for this by including a measure of their dose from this initiative in each MSA.

The dose variables for CJR and the Bundled Payments for Care Improvement initiative are defined as the baseline market share of providers that have ever been in the CJR model or the risk-bearing phase of the Bundled Payments for Care Improvement initiative, respectively. The baseline period for dose measurement was from October 2009 through September 2012, the three years before the first Bundled Payments for Care Improvement intervention period.[29]

We defined two CJR intervention periods:

- **CJR intervention period 1 (July 2015 – March 2016)** begins the quarter that the CJR model was announced but before it was implemented.
- **CJR intervention period 2 (April 2016 – December 2017)** begins the quarter that the CJR model took effect (April 1, 2016) and ends with the last quarter of PY2.

We interact each of the intervention period variables with their respective baseline market share doses to measure each market's exposure to CJR activity during the intervention periods. A similar approach is taken to control for the presence of the Bundled Payments for Care Improvement

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

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

activity. We also control for market-level characteristics, market and quarterly fixed effects, and a market-specific linear time trend. See Appendix E for a full description of methodology.

c. Results

There was no statistically significant difference in the volume of elective LEJR discharges between CJR MSAs and control MSAs. The estimated change in the discharge rate for elective LEJR due to the CJR model was a decrease of 0.033 per 1,000 FFS Medicare beneficiaries (p=0.42, Exhibit 29).

Exhibit 29: The CJR model did not have a statistically significant impact on the volume of elective LEJR discharges

Period	Predicted CJR MSA discharge rate (per 1,000 FFS beneficiaries)	Predicted counter-factual discharge rate (per 1,000 FFS beneficiaries)	Difference in discharge rates (per 1,000 FFS beneficiaries)	p-value
CJR intervention period 1 (July 2015 – March 2016)	3.220	3.225	-0.005	0.88
CJR intervention period 2 (April 2016 – December 2017)	3.295	3.340	-0.044	0.34
CJR intervention period 1 and 2 (July 2015 – December 2017)	3.273	3.305	-0.033	0.42

Source: Lewin analysis of Medicare claims and enrollment data for discharges from October 2007 through December 2017

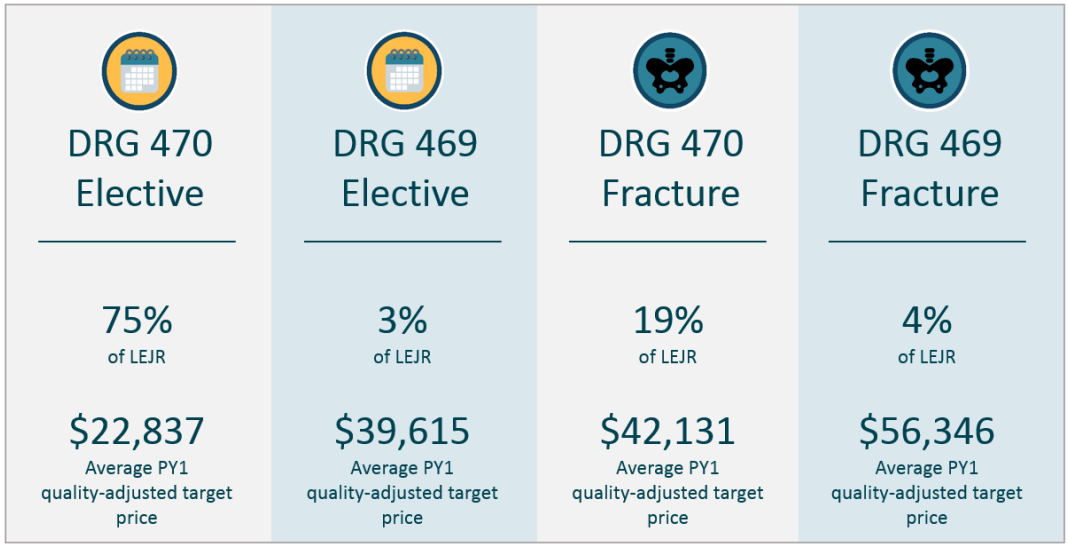
Notes: R-squared = 0.829.

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

2. Are there any indications that the CJR patient population was healthier in the intervention period than in the baseline period?

The ability of CJR participant hospitals to change their patient mix is more limited under the CJR model than under other bundled payment models because all hospitals within the MSA are participating in the model and have the same incentives. Further, under the CJR model, target prices differ by four episode groups determined by fracture status (elective or fracture) and MS-DRG (470 or 469) (Exhibit 30). As a result, a CJR participant hospital would need to change its mix of patients within an episode group to affect costs or quality outcomes. If a hospital increased the proportion of lower cost patients within one of the case-mix categories, it would be more likely to achieve average episode payments below its quality-adjusted target prices and earn reconciliation payments (see Section II.C for a discussion of changes in case-mix and reconciliation payments).

Exhibit 30: Under the CJR model, quality-adjusted target prices differ by presence or absence of a fracture and Medicare Severity-Diagnosis Related Group



Source: Lewin analysis of Medicare claims and enrollment data for episodes initiated in 2012 through 2014 that ended between April 2012 and March 2015 (baseline) and CJR payment contractor data for CJR participant hospitals in PY1.

Notes: DRG = diagnosis related group, LEJR = lower extremity joint replacement.

Changing the mix of patients within an episode group could be achieved if a hospital engaged in *upcoding* the sickest patients from MS-DRG 470 (without major complications and comorbidities) to MS-DRG 469 (with major complications and comorbidities). If the hospital more rigorously documented conditions that could qualify as major complications or comorbidities on the Medicare claim to increase patients coded as MS-DRG 469 instead of 470, then the patient complexity of both groups would decrease while the target prices remained the same. A hospital could also engage in *patient selection* by encouraging healthier beneficiaries to have an LEJR surgery or discouraging sicker Medicare beneficiaries from having an LEJR surgery. These activities would decrease patient complexity. They could also increase the volume of LEJR surgery and thus Medicare spending, or reduce access to LEJRs.

a. Key findings

- For elective MS-DRG 470, the least complex episode group, we found some evidence that the CJR patient population was relatively healthier in the intervention period than in the baseline period.
- For the more complex episode groups—elective MS-DRG 469, fracture MS-DRG 470, and fracture MS-DRG 469—we find no consistent evidence of changes in the severity of the CJR patient population.

b. Methods

We evaluated patient characteristics within each of the four episode groups to determine if CJR patients were on average healthier during the intervention period than they were historically. We examined changes in age, sex, race, Medicaid eligibility, disability status, health status, and prior health care use of LEJR patients from the baseline to the intervention period for CJR patients relative to control patients. Note that the impact analysis on payment, use, quality, and pain and functional status presented above controls for changes in these patient characteristics.

c. Results

For elective LEJRs in MS-DRG 470, there were changes in patient characteristics indicating that the CJR patient population was healthier in the intervention period than in the baseline period (Exhibit 31).

Disability. There was a statistically significant 0.8 percentage point (pp) decrease in the proportion of patients with a disability ($p<0.05$).[30]

Prior health care use. There were statistically significant decreases in the prevalence of prior health care use of 0.4pp for acute care hospital (ACH) stays ($p<0.05$) and 0.8pp for any institutional stays ($p<0.05$).

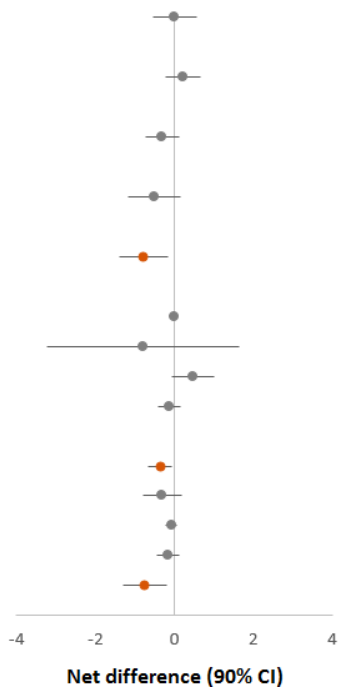
Moreover, while not statistically significant, there were decreases in the number of patients with other characteristics that are associated with higher episode costs (black, eligible for Medicaid, treated for smoking or obesity, HHA use, IRF stay, or SNF stay in the six months prior to the anchor hospitalization).

At the same time, there was a statistically significant decrease in the youngest age group, patients aged 20–64 ($p<0.10$), a population that is associated with lower episode costs even though they are eligible for Medicare because of disability (Appendix L).

[30] Patients with a disability are defined by the original reason for Medicare eligibility rather than the current reason for Medicare eligibility.

Exhibit 31: For elective LEJRs in MS-DRG 470, some indications of a healthier population

MS-DRG 470 elective	Domain	Net difference (90%CI)	*p*- value
Age	80+	0.0 (-0.5 to 0.6)	0.97
Sex	Female	0.2 (-0.2 to 0.7)	0.44
Race	Black	-0.3 (-0.7 to 0.1)	0.25
Medicaid	Eligible for Medicaid	-0.5 (-1.2 to 0.2)	0.22
Disability	**Disability, no ESRD**	**-0.8 (-1.4 to -0.2)**	**0.03**
Health status	HCC score	0.0 (-0.0 to 0.0)	0.77
	Obesity	-0.8 (-3.2 to 1.6)	0.60
	Diabetes	0.5 (-0.1 to 1.0)	0.15
	Tobacco use	-0.1 (-0.4 to 0.2)	0.47
Prior use	**ACH stay**	**-0.4 (-0.7 to -0.1)**	**0.05**
	HH use	-0.3 (-0.8 to 0.2)	0.30
	IRF stay	-0.1 (-0.2 to 0.1)	0.35
	SNF stay	-0.2 (-0.4 to 0.1)	0.37
	Any institutional stay	**-0.8 (-1.3 to -0.2)**	**0.02**



Source: Lewin analysis of Medicare claims and enrollment data for episodes initiated in 2012 through 2014 that ended between April 2012 and March 2015 (baseline) and episodes initiated during or after April 2016 that ended by December 2017 (intervention).

***Notes*:** Net differences that are significant at the 99%, 95%, or 90% significance level are indicated by dark, medium, and light orange shaded points, respectively. Rows with statistically significant characteristics are in bold.

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

For elective LEJRs in MS-DRG 469, there were no consistent changes in patient characteristics to indicate that the CJR patient population was healthier in the intervention period than in the baseline period (Exhibit 32).

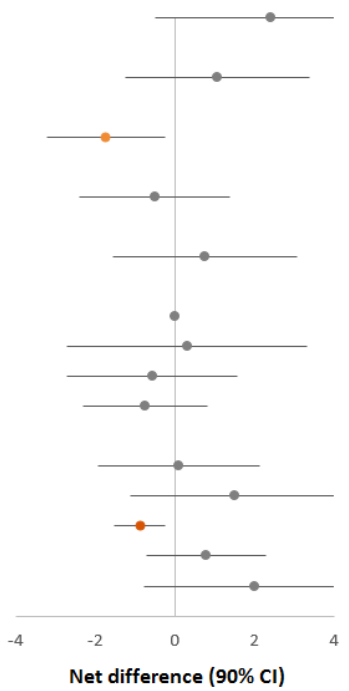
Race. There was a statistically significant 1.7pp decrease in the proportion of patients who were black ($p<0.10$), a characteristic associated with higher episode spending.

Prior health care use. There was also a statistically significant 0.9pp decrease in the proportion of patients who had an IRF stay in the six months prior to the anchor hospitalization ($p<0.05$). However, the remaining changes in measures of prior care use were not statistically significant and their direction pointed towards an increase in the prevalence of prior health care use.

Changes in the remaining characteristics were not statistically significant. Their direction was also mixed, and could be associated with lower or higher episode payments.

Exhibit 32: For elective LEJRs in MS-DRG 469, no consistent evidence to indicate a healthier population

MS-DRG 469 elective	Domain	Net difference (90% CI)	*p*- value
Age	80+	2.4 (-0.5 to 5.3)	0.17
Sex	Female	1.1 (-1.2 to 3.4)	0.44
Race	**Black**	**-1.7 (-3.2 to -0.2)**	**0.05**
Medicaid	Eligible for Medicaid	-0.5 (-2.4 to 1.4)	0.67
Disability	Disability, no ESRD	0.8 (-1.6 to 3.1)	0.59
Health status	HCC score	0.0 (-0.1 to 0.1)	0.88
	Obesity	0.3 (-2.7 to 3.3)	0.86
	Diabetes	-0.6 (-2.7 to 1.6)	0.67
	Tobacco use	-0.7 (-2.3 to 0.8)	0.43
Prior use	ACH stay	0.1 (-1.9 to 2.1)	0.92
	HH use	1.5 (-1.1 to 4.1)	0.34
	IRF stay	**-0.9 (-1.5 to -0.2)**	**0.03**
	SNF stay	0.8 (-0.7 to 2.3)	0.39
	Any institutional stay	2.0 (-0.8 to 4.8)	0.24



Source: Lewin analysis of Medicare claims and enrollment data for episodes initiated in 2012 through 2014 that ended between April 2012 and March 2015 (baseline) and episodes initiated during or after April 2016 that ended by December 2017 (intervention).

***Notes*:** Net differences that are significant at the 99%, 95%, or 90% significance level are indicated by dark, medium, and light orange shaded points, respectively. Rows with statistically significant characteristics are in bold.

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

d. Conclusion

These results suggest that the CJR patient population for elective MS-DRG 470 was relatively healthier in the intervention than in the baseline while there was no consistent change in the health of the CJR patient population for elective MS-DRG 469. Taken together, it does not appear that hospitals engaged in upcoding the sickest patients from MS-DRG 470 to MS-DRG 469. Recall from Section II.G.1, we found no evidence that volume changed because of the CJR model. Furthermore, the ratio of elective MS-DRG 470 to elective MS-DRG 469 episodes did not change because of the CJR model. Therefore, the increase in healthier patients for elective MS-DRG 470 may have been offset by a decrease in relatively more complex patients in MS-DRG 470. This indicates that at least a portion of the decrease in episode payments may be due to a healthier mix of patients at CJR participant hospitals rather than care redesign.

There is little ability to engage in patient selection among fracture patients, and we find no consistent evidence that the CJR patient population for fracture MS-DRG 470 and fracture MS-DRG 469 episodes was healthier in the intervention period than in the baseline (Appendix L).

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

Payments for services provided after the episode would not be considered during the reconciliation process, when episode payments are assessed relative to participant hospitals' quality-adjusted target prices. Thus, deferring services until after the episode could make it easier for participant hospitals to reduce their episode payments below their quality-adjusted target price to achieve a reconciliation payment without affecting episode payments.

a. Key findings

- The CJR model had no statistically significant impact on payments for services provided during the 30 days following the episode.

b. Methods

The same DiD methods described in Section II.A.2 were used for the analysis in this section. Additional details about the methodology are included in Appendix E.

c. Results

During the first two performance years, the CJR model had no statistically significant impact on payments for services provided during the 30 days following the episode (DiD= -$18, p=0.25, Appendix H). This suggests that CJR participant hospitals did not postpone services until after the end of the episode.

H. How did the impact of the CJR model vary for particular subgroups of MSAs, hospitals, and beneficiaries?

The design of the CJR model ensures that it includes various types of hospitals with diverse characteristics. This allows for a more robust, generalizable evaluation. It also allows us to examine questions about the ability of participant hospitals that face different circumstances to respond to the model, for example, hospitals with fewer resources to implement care redesign or those that have a more complex patient population.

We estimated the impact of the CJR model on the following subgroups based on MSA, hospital, and episode characteristics:

High vs. low-payment MSAs. CMS oversampled higher payment MSAs to participate in the CJR model because there may be greater opportunities to reduce payments in higher payment markets.

Additionally, hospitals in high-payment MSAs may be under more financial pressure to reduce payments because most started the CJR model with episode payments above their quality-adjusted target price (74%).

This subgroup category is defined based on MSA sampling strata used by CMS to select MSAs for participation in the CJR model. High-payment MSAs are those in strata 3, 4, 7, and 8. Low-payment MSAs are those in strata 1, 2, 5, and 6.

Hospital LEJR volume. Low-volume hospitals may be disadvantaged under the model if they have fewer resources with which to take action. Particularly high cost cases may also have a bigger impact on episode payment and quality for low-volume hospitals.

This subgroup category is defined based on a 3-year average at baseline (2012-2014): more than 200 LEJR, 100-200 LEJR, or less than 100 LEJR.

Patient complexity. It could be more difficult to reduce payments for more complex patients, such as patients with multiple or serious chronic conditions. HCC score, which accounts for the beneficiary's demographic characteristics and diagnoses, is an overall measure of patient complexity. An HCC score of 1.0 corresponds to average expected expenditure; higher HCC scores are associated with higher expected expenditures.

This subgroup category is defined based on quartiles of patient HCC scores at the start of the episode to examine the impact of CJR on patient subgroups with a range of patient complexity: less than 0.48 (Q1), 0.48 to less than 0.75 (Q2), 0.75 to less than 1.21 (Q3), and 1.21 or greater (Q4).

Fracture vs. elective episodes. It could be more difficult to reduce payments for more complex episodes and those for which the patient cannot be prepared in advance of the hospitalization, such as those with a fracture. Indeed, CMS provided different target prices for episodes due to fracture to account for the typically greater health care needs of these patients and wider variation in spending.[31] CMS also believes, however, that beneficiaries with hip fracture have the potential to benefit substantially from improved care coordination incentivized by the CJR model.[32] Because of both the concern and the promise for the treatment of fracture LEJR episodes under the CJR model, we analyzed the impact of the CJR model on these episodes for all claims-, assessment-, and patient survey-based outcomes.

Fracture episodes are identified using International Statistical Classification of Diseases and Related Health Problems (ICD) codes.

---

[31] Medicare Program Comprehensive Care for Joint Replacement Payment Model for Acute Care Hospitals Furnishing Lower Extremity Joint Replacement Services; A Final Rule by the Centers for Medicare & Medicaid Services, 80 FR 73273 (November 24, 2015) (codified at 42 CFR 510), p. 29.

[32] Ibid, p. 29.

Hospitals that we expect would not have volunteered for the CJR model had it been a voluntary initiative. We also analyzed patient survey outcomes for hospitals that we expect would not have volunteered for the CJR model had it been a voluntary initiative, which we identified as those with fewer than 68 LEJR discharges, based on our analysis of hospitals that volunteered for the Bundled Payments for Care Improvement initiative (Appendix E). These hospitals may be particularly disadvantaged under the model and participation could lead to large or significant differences in satisfaction with recovery and care management or care transitions between CJR and control respondents.

1. Key findings

- For all MSA, hospital, and episode subgroups examined, the CJR model resulted in decreases in average episode payments, and there were no statistically significant differences within subgroup categories, with the exception of the HCC category in which payment decreases for the subgroup with the lowest HCCs were less than payment decreases for the subgroups with higher HCCs.
- For fracture episodes, shifts from more intensive PAC settings and a reduction in readmission payments drove the decrease in average episode payments.
- For fracture patients discharged to SNF, there is some evidence that CJR resulted in lower rates of functional improvement, but by the end of the episode of care, CJR and control survey respondents had similar self-reported changes in functional status.
- For fracture episodes there is some evidence that CJR resulted in less satisfaction with care management and more reliance on caregivers.

2. Methods

We used the same DiD approach on claims and assessment data as described in Sections II.A.2 and II.F.2.a and the same approach to analyze patient survey data to estimate the impact of the CJR model on outcomes for subgroups of the CJR population.

For a more detailed description of the subgroup analysis methods, see Appendix E.

3. Results

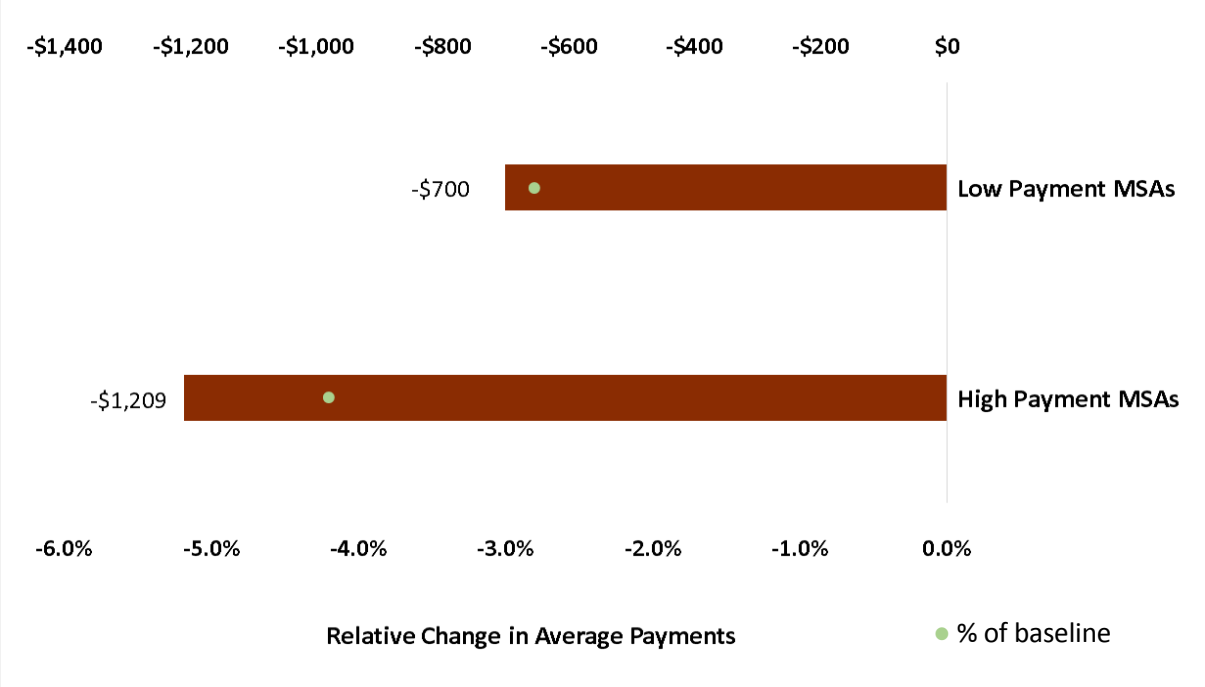
a. Variation in the impact of the CJR model on average episode payments for MSA, hospital, and patient subgroups

The CJR model resulted in statistically significant relative reductions in average episode payments for LEJR episodes across all subgroups.

High vs. low-payment MSAs. In high-payment MSAs, average episode payments decreased by $1,209 more for CJR episodes than for control group episodes from the baseline to the intervention period (4.2% of the CJR baseline, p<0.01, Exhibit 33). In low-payment MSAs, average episode payments decreased by $700 more for CJR episodes than for control group episodes from the baseline to the intervention period (2.8% of the CJR baseline, p<0.01).

Consistent with our hypothesis, the point estimates indicate that the high-payment MSAs decreased payments by more than the low-payment MSAs both in dollars and in percentage changes, however, the point estimates are not statistically different from one another (p=0.19).

Exhibit 33: CJR resulted in relative decreases in average episode payments for both high- and low-payment MSAs, PY1-2



Source: Lewin analysis of Medicare claims and enrollment data for episodes initiated in 2012 through 2014 that ended between April 2012 and March 1 2015 (baseline) and episodes initiated during or after April 2016 that ended by December 2017 (intervention).

Notes: Estimates that are significant at the 99%, 95%, or 90% significance level are indicated by dark, medium, and light orange shaded cells, respectively.

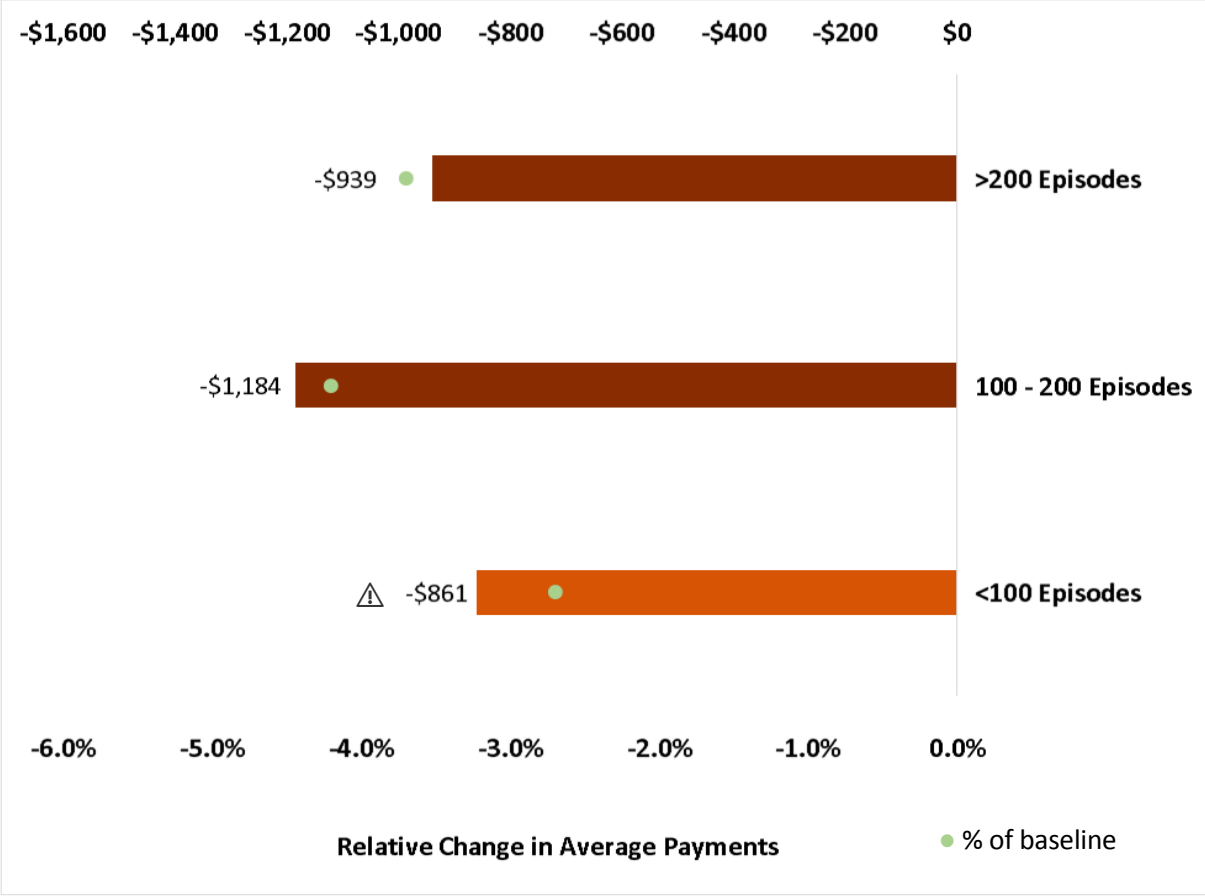
MSA = metropolitan statistical area.

Hospital LEJR volume. For high-volume hospitals (annual episodes >200), average episode payments decreased by $939 more for CJR episodes than for control group episodes (3.7% of the CJR baseline, $p<0.01$, Exhibit 34). For medium-volume hospitals (100-200 annual episodes), average episode payments decreased by $1,184 more for CJR episodes than for control group episodes (4.2% of the CJR baseline, $p<0.01$). For low-volume hospitals (annual episodes <100), average episode payments decreased by $861 more for CJR episodes than for control group episodes (2.7% of the CJR baseline, $p<0.05$).

Consistent with our hypothesis, the point estimates indicate that low-volume hospitals did not decrease payments by as much as medium- and high-volume hospitals both in dollars and in percentage terms, however, the point estimates are not statistically different from one another.[33] Furthermore, the low-volume subgroup failed the parallel trends test, which means that these results should be interpreted with caution. The control group for low-volume hospitals may not be a good representation of what would have happened to the low-volume CJR hospitals in the absence of the CJR model.

[33] The changes for low-volume hospitals (annual episodes <100) and medium-volume hospitals (100-200 annual episodes) are not statistically different ($p=0.38$). The changes for low-volume hospitals and high-volume hospitals (annual episodes >200) are not statistically different ($p=0.83$). The changes for medium-volume hospitals and high-volume hospitals are not statistically different ($p=0.34$).

Exhibit 34: CJR resulted in relative decreases in average episode payments for each LEJR volume subgroup, PY1-2



Source: Lewin analysis of Medicare claims and enrollment data for episodes initiated in 2012 through 2014 that ended between April 2012 and March 1 2015 (baseline) and episodes initiated during or after April 2016 that ended by December 2017 (intervention).

Notes: ⚠ Results need to be interpreted with caution because the measure fails the parallel trends test.

Estimates that are significant at the 99%, 95%, or 90% significance level are indicated by dark, medium, and light orange shaded cells, respectively.

Patient Complexity. For the least complex patients, the first HCC quartile, average episode payments decreased by $650 more for CJR episodes than for control group episodes (3.2% of the CJR baseline, $p<0.01$, Exhibit 35); for the second quartile, this decrease was $969 (4.2% of CJR baseline, $p<0.01$); for the third quartile, it was $1,109 (4.1% of the CJR baseline, $p<0.01$); and for the fourth quartile, the most complex patients, it was $1,330 (3.7% of the CJR baseline, $p<0.01$). Relative to the first quartile, the reduction for the second quartile was $319 more ($p<0.05$), the reduction for the third quartile was $459 more ($p<0.05$), and the reduction for the fourth quartile was $680 more ($p<0.01$).

Average episode payments decreased for patients across a range of complexity but, contrary to our hypothesis, the payment decrease for the least complex patient group (first quartile) is statistically significantly smaller than the payment decrease for the other patient groups (second through fourth

quartiles), suggesting that it is more difficult to reduce costs for the least complex patients, who already had relatively low episode payments in the baseline.

Exhibit 35: CJR resulted in relative decreases in average episode payments for each patient complexity subgroup, PY1-2

Source: Lewin analysis of Medicare claims and enrollment data for episodes initiated in 2012 through 2014 that ended between April 2012 and March 1 2015 (baseline) and episodes initiated during or after April 2016 that ended by December 2017 (intervention).

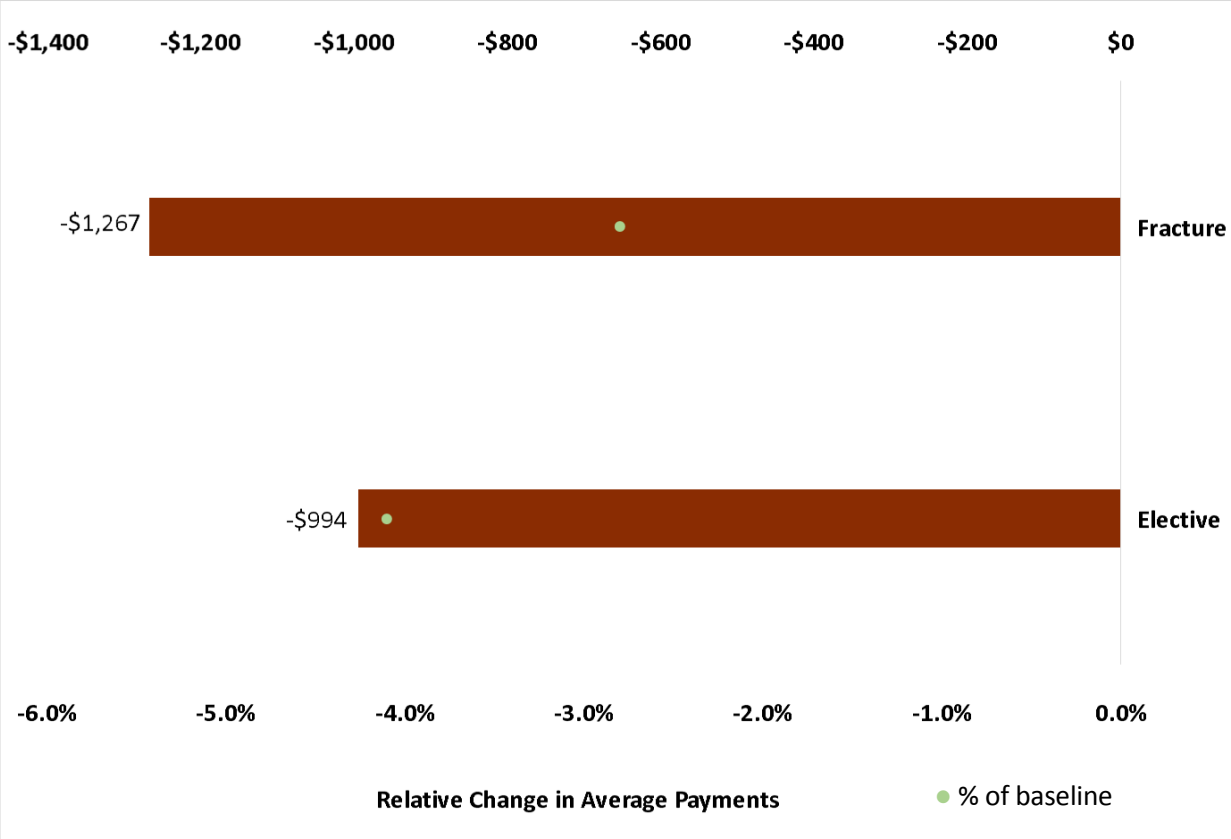
Notes: Estimates that are significant at the 99%, 95%, or 90% significance level are indicated by dark, medium, and light orange shaded cells, respectively.

Fracture vs. elective episodes. For fracture episodes, average episode payments decreased by $1,267 more for CJR episodes than for control group episodes from the baseline to the intervention period (2.8% of the CJR baseline, p<0.01). For elective episodes, average episode payments decreased by $994 more for CJR episodes than for control group episodes from the baseline to the intervention period (4.1% of the CJR baseline, p<0.01).

Contrary to our hypothesis that it would be more difficult to reduce payments for the more complex, unscheduled patients with LEJR due to fracture, average episode payments for fracture

episodes decreased both in dollars and in percentage terms (Exhibit 36). The percentage change decrease for fracture episodes is less than the percentage change decrease for elective episodes, but that difference is not statistically significant (p=0.49).

Exhibit 36: CJR resulted in relative decreases in average episode payments for fracture and elective episodes, PY1-2



Source: Lewin analysis of Medicare claims and enrollment data for episodes initiated in 2012 through 2014 that ended between April 2012 and March 1 2015 (baseline) and episodes initiated during or after April 2016 that ended by December 2017 (intervention).

Notes: The estimates in this exhibit are the result of a difference-in-differences (DiD) model. Estimates that are significant at the 99%, 95%, or 90% significance level are indicated by dark, medium, and light orange shaded cells, respectively.

b. Variation in the impact of the CJR model for fracture episodes

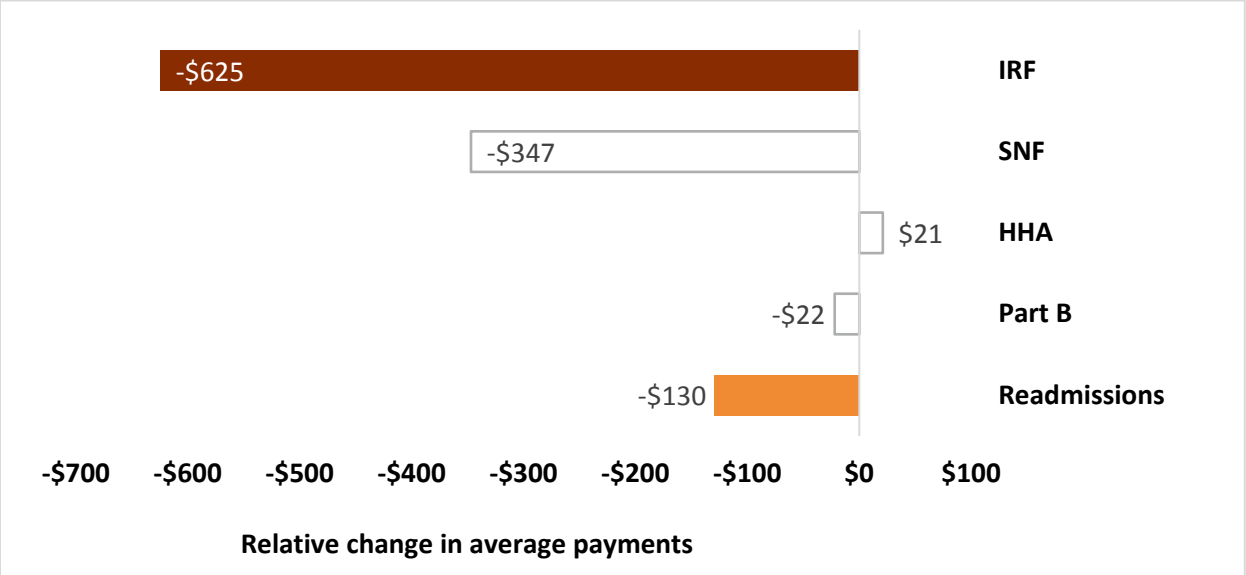
Compared to elective episodes, fracture episodes, on average, have much higher payments, greater use of institutional PAC, and greater rates of unplanned readmissions, emergency department use, and mortality. For these complex episodes, this section reports the results for all claims-, assessment-, and patient survey-based outcomes. The results for elective episodes are reported in Appendix H.

Average episode payments, service-level payments, and use

For fracture episodes, average episode payments decreased by $1,267 more for CJR episodes than for control group episodes from the baseline to the intervention period (p<0.01), which equates to a 2.8% decrease from the CJR baseline. The relative decrease in average episode payments was the

result of relative reductions in IRF and readmission payments (Exhibit 37). Average IRF payments for fracture episodes decreased by $625 more for CJR episodes than for control group episodes from baseline to intervention (p<0.01), a 14.7% decrease from the CJR baseline. Average readmission payments decreased by $130 more for CJR fracture episodes (p<0.10), a 5.5% decrease from the CJR baseline.

Exhibit 37: Inpatient rehabilitation facility and readmission payments significantly decreased for fracture episodes, PY1-2



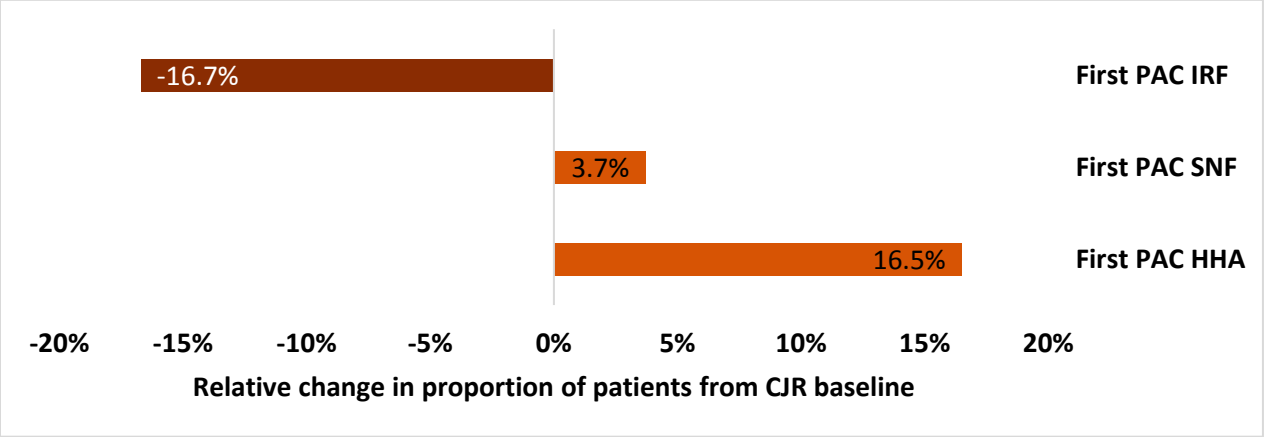
Source: Lewin analysis of Medicare claims and enrollment data for episodes initiated in 2012 through 2014 that ended between April 2012 and March 2015 (baseline) and episodes initiated during or after April 2016 that ended by December 2017 (intervention).

Notes: The estimates in this exhibit are the result of a difference-in-differences (DiD) model. DiD estimates that are significant at the 99%, 95%, or 90% significance level are indicated by dark, medium, and light orange shaded bars, respectively.

HHA = home health agency, IRF = inpatient rehabilitation facility, PY = performance year, SNF = skilled nursing facility.

The decrease in average IRF payments is explained by the relative reduction in fracture episodes first discharged to an IRF after the hospitalization (Exhibit 38). The percent of fracture episodes first discharged to IRFs decreased more for CJR than for the control group from baseline to intervention, equating to a 16.7% decrease from the CJR baseline (p<0.01).

Exhibit 38: Fewer CJR fracture patients were discharged to inpatient rehabilitation facilities, more discharged to skilled nursing facilities or home health agencies, PY1-2



Source: Lewin analysis of Medicare claims and enrollment data for episodes initiated in 2012 through 2014 that ended between April 2012 and March 2015 (baseline) and episodes initiated during or after April 2016 that ended by December 2017 (intervention).

Notes: The estimates in this exhibit are the result of a difference-in-differences (DiD) model. DiD estimates that are significant at the 99%, 95%, or 90% significance level are indicated by dark, medium, and light orange shaded bars, respectively.

HHA = home health agency, IRF = inpatient rehabilitation facility, PAC = post-acute care, PY = performance year, SNF = skilled nursing facility.

In parallel, there were relative increases in the percent of fracture episodes first discharged to SNF (3.7% increase from the CJR baseline; $p<0.05$) and to HHA (16.5% increase from the CJR baseline; $p<0.05$) settings (Exhibit 38). Similar to the results for all LEJR episodes, these changes indicate CJR participant hospitals responded to the model by shifting care from more expensive to less expensive PAC settings for fracture episodes.

There was no statistically significant change in SNF payments because the relative increase in the percent of fracture episodes first discharged to SNF was offset by a statistically significant relative decrease in the number of SNF days. Among fracture episodes with a SNF stay, the number of SNF days decreased by 1.7 days more for CJR episodes than for control group episodes from baseline to intervention ($p<0.01$) (Appendix H).

There was also no statistically significant change in HHA payments, despite the substantial increase in those first discharged to HHA, because the proportion of fracture episodes with HHA care at *any* time during the episode did not experience a statistically significant change.

Quality of care

For fracture episodes, despite significant shifts in care from more expensive to less expensive PAC settings, quality of care, as measured by the unplanned readmission rate, emergency department use, and the mortality rate, was maintained (Appendix H).

Even though we observed no change in the unplanned readmission rate, readmission payments significantly decreased for fracture episodes. These payments include both planned and unplanned

readmission payments, and changes in average readmission payments could be due to changes in the readmission rate or the average readmission payment among those with a readmission. The primary driver for the decrease in readmission payments was a decrease in the unplanned average readmission payment among CJR fracture episodes with an unplanned readmission, indicating a decrease in the intensity of unplanned readmissions.

Functional status and pain

This section presents the results from the patient survey and the PAC assessments on functional status and pain for patients with an LEJR following a fracture. The survey data represent changes in functional status later in the course of recovery, whereas the PAC assessments represent changes made during the PAC stay, which typically ends earlier in the recovery period.

Patient survey results

In both the CJR and control groups, survey respondents with an LEJR following a fracture reported decreased functional status across all measures and more use of pain medication from a week before hospitalization (i.e., before their fracture) through the time of the survey (Exhibit 39). The decline in functional status and increased use of pain medications among fracture respondents may be because respondents were asked to recall their functional status the week prior to surgery, which was required due to a fracture. It is likely that patients with fractures had fewer functional limitations prior to their fracture than did those with elective LEJR surgeries had a week before their surgery and would thus be more likely to report declines in functional status.

Overall, the model did not have a statistically significant effect on patients' functional status or pain. CJR respondents indicated a statistically significantly lower decline in their ability to rise from sitting than did control respondents ($p<0.05$), but there were no other statistically significant differences between the two groups.

Exhibit 39: Changes in functional status and pain among survey respondents with LEJR due to a fracture were similar

Survey measure	Change in self-reported measure from before the hospitalization to after the episode *Higher value represents a more favorable change*		Difference between CJR and control groups
	CJR	Control group	
Ability to walk by yourself without resting[a]	-0.77	-0.76	-0.01
Difficulty walking up or down 12 stairs[b]	-0.56	-0.53	-0.02
Difficulty rising from sitting[a]	-0.23	-0.36	0.13
Difficulty standing[a]	-0.28	-0.29	0.02
Use of a mobility aid[c]	-0.66	-0.61	-0.04
Difficulty getting on/off the toilet[a]	-0.06	-0.10	0.04
Frequency that pain interferes with normal activities[a]	-0.34	-0.30	-0.04
Medication use for pain in the joint you had replaced[b]	-0.39	-0.33	-0.06

Source: Lewin analysis of patient survey data for episodes with discharge in March, April, September, and October 2017.

Notes: The change in a given measure refers to the difference between a respondent's self-reported status at the time of the survey and the respondent's recalled pre-hospital status. Estimated changes, and the difference between changes in the CJR and control group, are reported in "level" terms.

The estimates in this exhibit are the results of a cross-sectional regression model, weighted for sampling and nonresponse. Estimates that are significant at the 99%, 95%, or 90% significance level are indicated by dark, medium, and light orange shaded cells, respectively.

[a] Indicates the question has 5 possible responses (i.e., "levels"), and the change could range from -4 to 4.

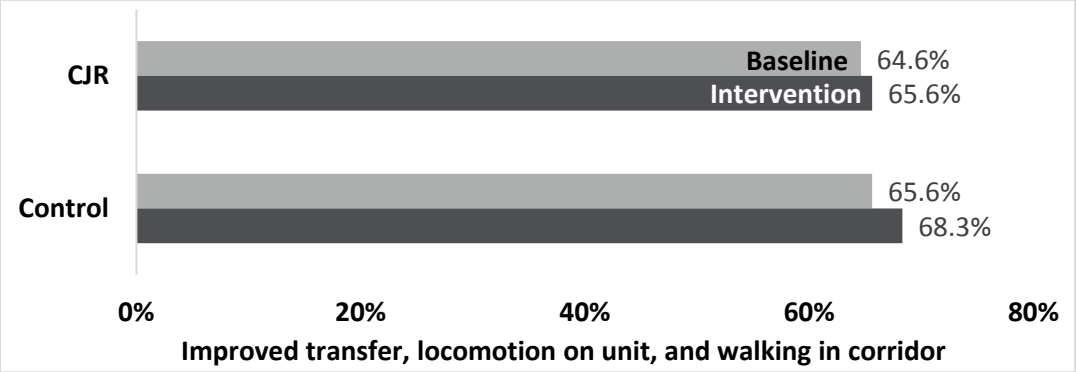
[b] Indicates the question has 4 possible responses (i.e., "levels"), and the change could range from -3 to 3.

[c] Indicates the question has 3 possible responses (i.e., "levels"), and the change could range from -2 to 2.

Assessment-based results

Almost 70% of CJR fracture patients were discharged to a SNF during the baseline period and their stays averaged 40 days. The majority of CJR patients with a fracture who were first discharged to a SNF improved in mobility, which was measured as transfer, locomotion on unit, and walking in corridor, during their SNF stay. The proportion of CJR patients with improvement in mobility increased by 1.0 percentage point from baseline to intervention (from 64.6% to 65.6%). For the control group, the proportion of patients with improvement in mobility increased by 2.7 percentage points (from 65.6% to 68.3%) (Exhibit 40). As a result, the proportion of CJR patients with improvement in mobility decreased by 1.7 percentage points relative to the control patients, or 2.6% from the CJR baseline period ($p<0.05$, Appendix H).

Exhibit 40: The proportion of fracture patients with improved mobility from baseline to intervention increased for CJR patients discharged to SNF, but increased more for control patients, PY1-2.

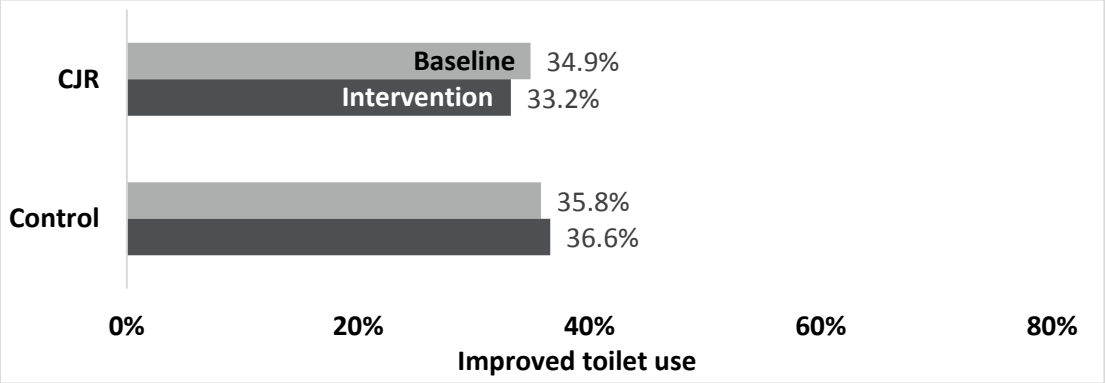


Source: Lewin analysis of Minimum Data Set (MDS) data for episodes initiated in April 2012 and March 2015 (baseline) and episodes initiated during or after April 2016 that ended by September 2017 (intervention).

Notes: PY = performance year, SNF = skilled nursing facility.

Only about one-third of CJR fracture patients first discharged to a SNF improved in toilet use during their SNF stay, and the proportion with improved toilet use decreased by 1.7 percentage points from baseline to intervention (from 34.9% to 33.2%); for the control group, it increased by 0.9 percentage points (from 35.8% to 36.6%) (Exhibit 41). As a result, the proportion of CJR fracture patients first discharged to a SNF who improved toilet use decreased by 2.5 percentage points relative to the control group, or 7.3% from the CJR baseline period ($p<0.05$, Appendix H).

Exhibit 41: The proportion of fracture patients with improved toilet use from baseline to intervention decreased for CJR patients discharged to SNF, but increased for control patients, PY1-2.



Source: Lewin analysis of Minimum Data Set (MDS) data for episodes initiated in April 2012 and March 2015 (baseline) and episodes initiated during or after April 2016 that ended by September 2017 (intervention).

Notes: PY = performance year, SNF = skilled nursing facility.

There was no statistical difference in the proportion of fracture patients without moderate or severe pain between CJR and control patients first discharged to a SNF from the baseline to the intervention (Appendix H). In addition, the CJR model did not have a statistically significant effect on changes in mobility scores among fracture patients who were first discharged to an IRF. Nor did

it have a statistically significant effect on the proportion of patients with improvements in functional status or pain among fracture patients first discharged to an HHA (Appendix H).

Sensitivity findings

We conducted additional analyses to better understand the changes in functional status of CJR fracture patients who were discharged to PAC relative to the control group. The CJR model resulted in relative reductions in the length of SNF stays among fracture patients (-1.7 days, $p<0.01$, Appendix H). This indicates that CJR fracture patients first discharged to a SNF had less time for improvements in functional status and pain while receiving SNF care. To understand whether less time in the SNF setting contributed to the relative decreases in the proportion of patients with improvement, we measured the average time between assessments for fracture patients discharged to a SNF, and we re-calculated the pain and functional status estimates controlling for the number of days between assessments.[34]

Consistent with the relative decrease in the number of SNF days, the number of days between patient assessments decreased by 2.0 days more for CJR fracture patients ($p<0.01$, Appendix J). After controlling for the number of days between patient assessments in the risk adjustment models, the estimated impact of the CJR model on toilet use for fracture patients decreased from a 7.3% decrease to a 5.2% decrease from the CJR baseline, and was no longer statistically significant. The estimated impact of the CJR model on motion for fracture patients decreased from a 2.6% decrease to a 1.1% decrease and was no longer statistically significant (Appendix J).

In another sensitivity analysis, we found that the majority of fracture patients with no improvement in their functional status during their PAC stay maintained their functional status, that is, they did not decline in their status (Appendix J).

Care experiences based on the patient survey

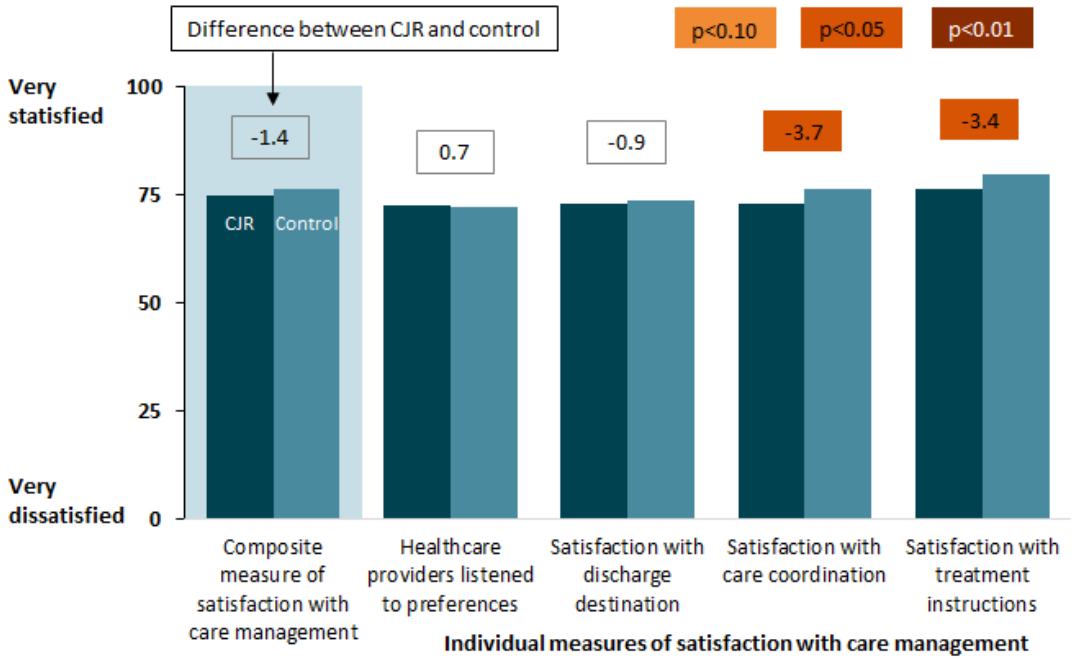
Satisfaction

Compared to the patient survey results for all LEJR episodes, respondents with a fracture were less satisfied with care management and with their overall recovery (Appendix K). CJR and control group respondents with fractures reported similar levels of satisfaction with overall physical recovery (Appendix K). Among respondents with fractures, CJR respondents reported slightly lower levels of satisfaction than control respondents on two individual measures of care management. CJR respondents scored 3.7 points (out of 100) lower in satisfaction with care coordination ($p<0.05$), and 3.4 points lower in satisfaction with treatment instructions received ($p<0.05$). Our clinical review panel indicated that the magnitude of the difference was neither concerning nor clinically relevant. Differences in the satisfaction with care management

[34] Because the CJR model may impact both the length of SNF stays and ADL outcomes, the number of days between assessments is not included as a causal risk factor in the risk adjustment models for the main analysis. However, including the variable in the risk adjustment models helps us determine whether the relative change in the rate that CJR patients improved in functional status can be explained by its association with the relative change in number of days between SNF assessments among CJR patients.

composite score, which summarized four care management measures, were not statistically significant, suggesting that overall, differences between CJR and control respondents were marginal (Exhibit 42).

Exhibit 42: CJR respondents with fractures were less satisfied with care coordination and treatment instructions



Source: Lewin analysis of patient survey data for episodes with discharge in March, April, September, and October 2017.

Notes: Differences that are significant at the 99%, 95%, or 90% significance level are indicated by dark, medium, and light orange shading, respectively.

All outcomes are scaled from 0 to 100 points, where 0 = very dissatisfied, 25 = somewhat dissatisfied, 50 = neutral, 75 = somewhat satisfied, and 100 = very satisfied.

The composite summarizes all four measures of satisfaction with care management.

Differences between CJR and control outcomes are reported in point terms.

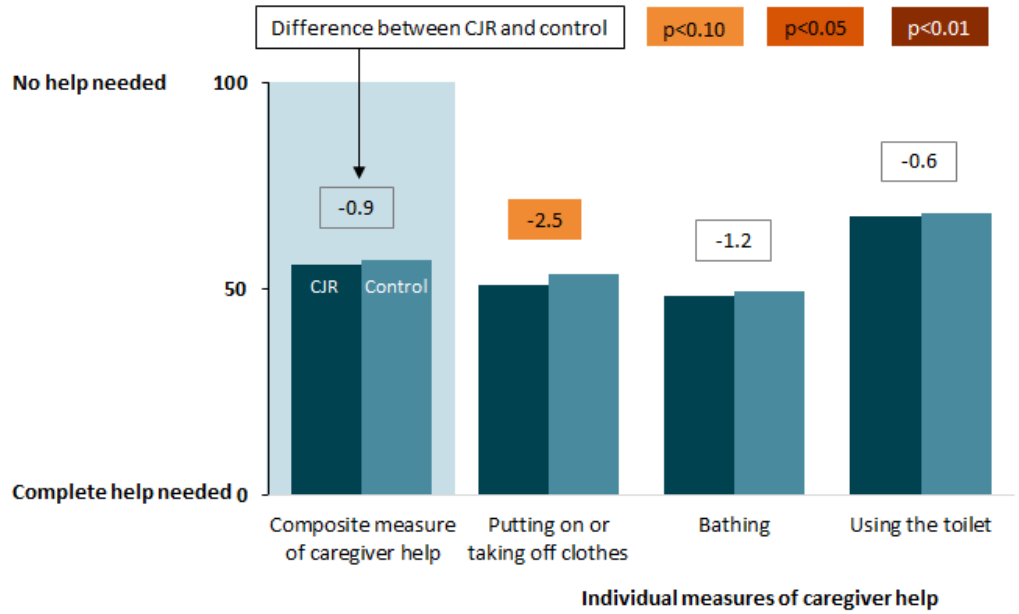
Care Transitions

There were no significant differences in the proportion of CJR and control fracture respondents reporting that they were discharged on time, had appropriate levels of post-discharge care, or had access to the durable medical equipment they needed (Appendix K).

Caregiver help

Approximately 95% of CJR and control fracture respondents received at least some help from a caregiver after returning home (Appendix K). Among those who received help, CJR respondents scored 2.5 points (out of 100) lower than control respondents on help needed putting on or taking off clothes ($p<0.10$, Exhibit 43). Differences in the composite score, which summarized help across three ADLs, were not statistically significant. Our clinical review panel indicated that that the difference was not concerning or clinically meaningful.

Exhibit 43: CJR respondents required slightly more help putting on or taking off clothes



Source: Lewin analysis of patient survey data for episodes with discharge in March, April, September, and October 2017.

Notes: Differences that are significant at the 99%, 95%, or 90% significance level are indicated by dark, medium, and light orange shading, respectively.

Respondents were only asked about the amount of help needed with a given activity of daily living if they indicated that they received caregiver help.

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

The composite summarizes the amount of help needed across all three activities of daily living.

Differences between CJR and control outcomes are reported in point terms.

Finally, we analyzed patient survey data for hospitals that we expect would not have volunteered for the CJR model had it been a voluntary initiative. We did not see any large or significant differences in functional status, pain, or satisfaction with care relative to the control group.

An In-Depth Look: Hospitals unlikely to participate in a voluntary episode-based payment model

We analyzed patient survey data for hospitals that we expect would not have volunteered for the CJR model had it been a voluntary initiative, which we identified as those with fewer than 68 LEJR discharges, based on our analysis of hospitals that volunteered for the Bundled Payments for Care Improvement initiative (Appendix E).

Survey responses were similar between CJR and control patients. We did not observe any large differences in functional status or pain between CJR and control respondents. We also did not observe any large or significant differences in satisfaction with recovery and care management or care transitions between CJR and control respondents. In fact, CJR respondents who received caregiver help at home indicated that they did not need as much assistance putting on or taking off clothes compared to control respondents. Complete survey results for patients discharged from hospitals with fewer than 68 LEJR discharges are reported in Appendix K.

c. Conclusion

The CJR model resulted in statistically significant relative reductions in average episode payments for LEJR episodes across a range of MSA, hospital, and patient subgroups. Average episode payments decreased in MSAs with historically high and historically low payments, for hospitals with high, medium and low LEJR episode volume, for patients with varying levels of complexity (based on HCC score), and for patients with planned LEJRs and those due to fracture. There were no statistically significant differences in the impact of the CJR model on average episode payments within subgroup categories with one exception. Payment decreases for the least complex patient subgroup were less than payment decreases for the most complex patient subgroups. This suggests that it is more difficult to reduce payments for patients with relatively low episode payments in the baseline.

For fracture episodes, shifts from more intensive PAC settings and a reduction in readmission payments drove the decrease in average episode payments. Despite significant shifts in care from more expensive to less expensive PAC settings, quality of care, as measured by the unplanned readmission rate, emergency department use, and the mortality rate, was maintained. The self-reported outcomes from the patient survey indicate that CJR and control patients with a fracture had similar levels of improvement in functional status from the week prior to the surgery until after the end of the episode. This indicates that longer-term functional status and pain outcomes for fracture patients were not adversely affected under the CJR model.

We also examined shorter-term changes in functional status and pain for fracture patients discharged to an IRF, SNF, or HHA. For fracture patients first discharged to an IRF or HHA, the CJR model did not have a statistically significant impact on functional status or pain. For those first discharged to a SNF, the assessment-based outcomes generally indicated a decrease in the proportion of CJR fracture patients who improved their functional status from baseline to intervention compared to the control group. Notably, the majority of fracture patients in both the

CJR and the control groups did not improve in toilet use and there was a decrease in the proportion of CJR fracture patients who improved from baseline to intervention. By contrast, there was an increase in the proportion of control group fracture patients whose functional status improved.

CJR fracture patients discharged from the hospital to a SNF had less time to improve their functional status while receiving PAC because of their shorter SNF stays. The decrease in the number of days between assessments explained some, but not all, of the relative decrease in the proportion of fracture patients with improvement in functional status in the SNF setting.

The SNF assessment results could reflect increased patient complexity as the CJR model shifts PAC services from more intense to less intense PAC settings. Although we risk adjusted our estimates to control for differences in patient mix, there could be changes in unobserved patient characteristics that we are not capturing in our models.

CJR survey respondents with fractures were less satisfied with care coordination and treatment instructions and reported a slightly higher reliance on caregivers than control survey respondents with fractures. We asked the clinical review panel members about the importance of the assessment-based and self-reported outcomes and they concluded that the statistically significant differences were not clinically meaningful.

I. What key factors and model features influenced hospitals' choice of response to the CJR model?

The CJR model is not prescriptive in what hospitals should do to reduce LEJR episode payments and improve quality of care. How and whether hospitals choose to respond to the CJR model are likely affected by their specific circumstances, like characteristics of the local market and availability of internal resources.

1. Key findings

Representatives from hospitals reported that:

- Responses to the CJR model were made in the context of other hospital and market priorities and their assessment of potential gains or losses
- The opportunity to prepare for future bundled payment models was a strong motivating factor in hospitals' responses to the CJR model
- Prior hospital initiatives or participation in other payment and delivery models helped them prepare for the CJR model

2. Methods

We synthesized information from site visit interviews, telephone interviews, and the hospital survey to understand the factors that influenced hospitals' responses to the model. More detail on our methods is in Appendix E, and hospital survey questions are in Appendix M.

3. Results

a. Perceptions of financial pressure and model incentives

Representatives from nearly every hospital we spoke with described the importance of their orthopedic (and LEJR) service line to hospital revenues; however, they varied widely in their views of the financial impact of the CJR model. Many noted that potential reconciliation payments from the CJR model were small in comparison to other hospital or system initiatives and "not big money for the hospital." Representatives from some of these hospitals explained that they have a stronger strategic focus on the orthopedic service line for commercial payers than for Medicare, due to the larger profit margin. Others described minimal financial pressure due to anticipated gains under the model. For example, one hospital interviewee stated that the hospital was positioned to financially benefit from the quality-adjusted target price because it was highly efficient with excellent quality in a high-payment region. Representatives of other hospitals with higher shares of Medicare patients described concerns with the ability to offset losses with gains on commercial cases and mentioned that the effort and expense required to prepare for the CJR model would likely not be offset by reconciliation payments.

From the Case Studies

Health system representatives from **Hospital A** felt that the CJR model was not a significant source of financial pressure, because the hospital bears financial risk for more patients under the Medicare Shared Savings Program (MSSP) ACO, which requires participants to be accountable for total health care spending for the assigned population. These representatives explained that participation in the CJR model was worthwhile because they perceived bundled payments to be the future of health care reimbursement.

Irrespective of their perception of financial pressure, many interviewees indicated that the opportunity to prepare for future bundled payment models was a strong motivating factor in their response to the model. Hospital interviewees often explained that other hospitals not participating in the CJR model may be at risk of "getting into the game late." Interviewees felt that as early participants, they may have an advantage over late adopters. If CMS expands the model to non-participating hospitals, regional costs may have decreased substantially, potentially creating an even larger difference between high cost hospitals and regional averages.

b. Opportunity analysis

Representatives from most hospitals reported assessing how much the hospital could gain or lose under the CJR model and whether their potential response would be "worth the effort." Some reported contracting with external vendors to conduct a cost-benefit analysis, while others

indicated that this analysis was conducted internally by hospital staff or with the support of the hospital system.

Hospital interviewees indicated that the Medicare data they received as participants in the CJR model was useful in determining how to respond to the model. Many indicated that prior to the CJR model they did not have data on the entire episode and noted its value in understanding total episode costs, including the contribution of PAC use, especially SNF length of stay, and hospital readmissions. Frequently, they indicated that PAC utilization represented the largest opportunity to reduce episode payments. Some also noted that the opportunity analysis identified potential internal cost saving strategies, which were subsequent areas of focus for hospital actions.

An In-Depth Look: Total Arthroscopic Knee Replacement

On December 14, 2017, CMS indicated that beginning in 2018 Medicare would remove total knee arthroplasty (TKA) from the inpatient only list. This means that Medicare would begin paying for the procedure in the outpatient setting, although CMS has reiterated that this does not mean that TKA must be performed on an outpatient basis. Interviewees from most hospitals referenced this policy change during site visit discussions of key factors and model features.

Interviewees were concerned that by permitting TKA surgery in the outpatient setting, only the higher risk, and thus more expensive, patients would have their surgery in the inpatient setting. Interviewees indicated that fewer healthier patients with lower post-acute care spending would have inpatient surgeries, resulting in higher average CJR episode payments, which could jeopardize their financial success under the model. In the Final Rule implementing this change, CMS explained that although an increasing number of TKA cases may shift to the outpatient setting, it does not expect a large decrease in the volume of cases currently performed in the inpatient setting before the end of the CJR model in 2020. CMS also said it is monitoring outpatient TKAs to determine if it needs to change quality-adjusted target price calculations under the CJR model due to shifts to the outpatient setting.

(Please note that the analyses in this report are not affected by this policy change.)

c. Hospital resources

Hospital interviewees described internal resources and broader organizational factors that influenced hospital response to the CJR model. Many discussed the role of the hospital system, although there was substantial variation in the level of involvement and influence on hospital response to the model. Some hospitals were in systems with a centralized approach for all of their CJR participant hospitals. Other systems were minimally or not involved. Systems often provided data analytic services, which helped member hospitals use CMS data to understand episode payments during the baseline period.

Several hospital interviewees reported that their response was influenced by strong surgeon engagement in the model. One hospital that opted to remain in the CJR model after removing the mandatory requirement, did so because of strong surgeon engagement; the interviewee stated that

the surgeons "felt like what they were working toward was positive." Additionally, interviewees described the importance of a surgeon champion, often noting that a positive past experience with care redesign leadership supported the hospital's response to the CJR model.

d. Hospital experiences

Hospital survey results show that 60% of CJR participant hospital respondents had experience in value-based payment (VBP) models. At the time of the survey, 45% reported current participation in commercial payer VBP models, 35% reported participation in other Medicare VBP models, and 28% reported participation in Medicaid VBP models.

> *"We are surrounded by facilities that have some participation in BPCI. We have a couple of physicians that are familiar with BPCI and therefore bring some interesting ideas to the table about things that are done at different facilities."*
>
> *- System administrator*

Hospital representatives from the site visits and telephone interviews indicated that their prior hospital initiatives or participation in other VBP and delivery models helped prepare them for the CJR model. Hospital representatives reported leveraging LEJR care pathways developed for other episode-based payment approaches, such as Medicare's Bundled Payments for Care Improvement initiative or commercial payer bundles. Staff at hospitals with relevant prior experiences indicated more capacity than interviewees from other hospitals to identify areas for improvement and implement care redesign changes to succeed under the CJR model. Hospital representatives we contacted for telephone interviews indicated that Bundled Payments for Care Improvement participation in particular was especially helpful in setting up gainsharing agreements. Most of the health system administrators reported that another hospital in their system participated in the Bundled Payments for Care Improvement initiative for LEJR surgeries and provided tools, such as a template for a gainsharing agreement, for use by the CJR participant hospitals in the system. Some interviewees also stated that their CJR model-related activities benefitted from a "trickle down" or "halo" effect as the hospital prepares to participate in multiple episodes in Medicare's new Advanced Bundled Payments for Care Improvement initiative. Experience engaging surgeons in the CJR model, establishing gainsharing agreements, and other changes may spill over to other episodes and to engagement with other physician groups under the new initiative.

A number of interviewees reported that surgeons with past VBP model experience were more active in the hospital's response to the CJR model. They said that these surgeons brought interesting ideas about care redesign and were more willing to change practice patterns than those without prior experience. Specifically, hospitals reported that surgeons who had Bundled Payments for Care Improvement initiative experience were more likely to be interested in gainsharing for the CJR model, especially if they were previously financially successful in that initiative.

e. Conclusion

Generally, hospital representatives reported that they made decisions about their response to the model in the context of the hospital's market, complete orthopedic service line, resources, and

experience. Often, the influence of the CJR model was not distinguishable from market factors that influenced decisions about the overall orthopedic service line.

J. What did CJR participant hospitals do to redesign care for their LEJR patients?

Hospitals participating in the CJR model are expected to implement care redesign during the pre-surgical, inpatient, or post-discharge periods with the goals of reducing total episode payments or improve quality of care or both.

1. Key findings

Hospital representatives reported:

- Increasing physical therapy prior to surgery and identifying higher risk patients early to facilitate safe discharge home and optimize patient outcomes
- Changing pain management and physical therapy protocols
- Extending patient follow-up and developing PAC protocols and preferred PAC provider networks to discharge more patients directly home, improve overall quality, and prevent readmissions

2. Methods

We synthesized information from site visit interviews, telephone interviews, claims analysis, the CRP, and the hospital survey to understand experience with enhanced or new initiatives implemented for LEJR patients. For more detailed information about these methods, please see Appendix E**.**

3. Results

a. Prior to admission

Consistent with our impact estimates of lower SNF payments (Section II.D.3), hospital respondents indicated that reducing SNF care is among the key objectives of their redesign activities. Hospital representatives reported initiating discharge planning well before the hospital admission to educate patients about the most appropriate hospital discharge destination for them and to identify high-risk patients to optimize health outcomes, which would reduce their need for institutional post-acute care.

Patient education

Many hospital representatives described pre-surgical patient education classes as an important element to their response to the CJR model. The classes were used to set patient and caregiver expectations for discharge destination, continue discharge planning, identify and mitigate risks to successful recovery, and build caregiver engagement. Several hospitals described specific efforts to enhance their pre-surgical patient education activities because of CJR, which are described in the case studies.

From the Case Studies

In response to the CJR model, **Hospital E** made pre-surgical education class "semi-mandatory" and added class attendance to the surgeons' evaluation metrics in the co-management agreement. Interviewees indicated that patients who attended the pre-surgical class had better surgical outcomes and higher levels of satisfaction.

Similarly, most hospital survey respondents indicated that they were implementing patient education prior to admission (80%), and half of these respondents indicated that the CJR model influenced their decision to implement or further enhance patient education activities.

Risk stratification and patient optimization

According to the hospital survey, 75% of respondents implemented standardized patient assessments of environmental factors influencing patient recovery prior to scheduling surgery and 68% implemented specialized care plans based on patient risk stratification. Roughly two-thirds of respondents indicated the CJR model influenced these implementation decisions.

From the Case Studies

Representatives from **Hospital G** described their experience implementing new risk stratification and patient optimization strategies as a result of the CJR model. In the first 6 months of implementation, the hospital identified 19 out of 67 patients who were newly diagnosed with or had out-of-control diabetes. Patient optimization efforts, including referrals to endocrinology and day of surgery blood glucose tests, resulted in 12 of those patients continuing to surgery. The interviewees indicated that the CJR model educates physicians to optimize patients to be in the best condition possible before a procedure.

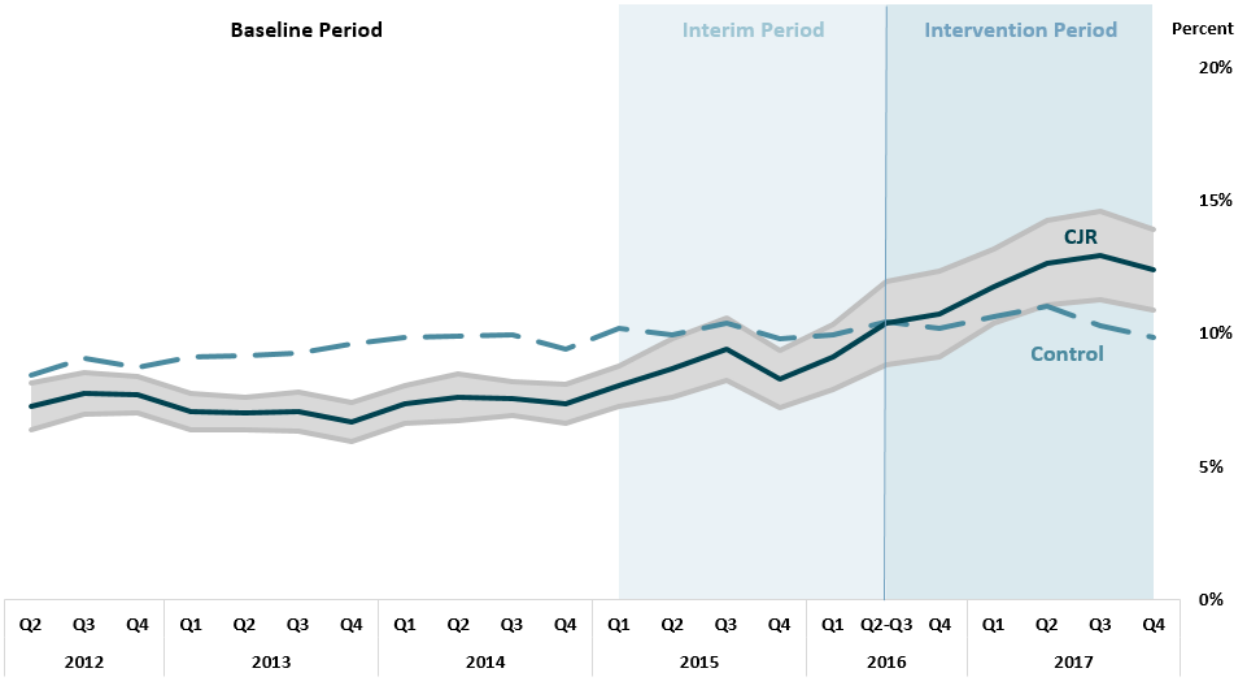
Hospital interviewees indicated that they used risk stratification protocols to identify higher risk patients prior to admission so that they could mitigate barriers to safe discharge home or address modifiable risk factors to optimize patient outcomes. Some hospitals, for example, used the Risk Assessment and Prediction Tool (RAPT), a standardized pre-surgical survey, to predict the best discharge destination for a patient. Others created tools to assess medical risk factors and environment and social considerations, like the availability of caregiver support. Interviewees also noted new emphasis on pre-surgical patient optimization to address modifiable risk factors such as weight, control of diabetes, or tobacco use. Some interviewees discussed "hard stops" for these factors, or imposing thresholds for health metrics before a patient would be scheduled for surgery. Many surgeons discussed implementing at least some sort of guideline or target range for these metrics.

Prehabilitation

Prehabilitation or "prehab" is physical therapy prior to elective LEJR surgery. It is becoming more common under the CJR model according to our site visit interviews, our clinical expert panel, and Medicare claims data. The proportion of patients receiving prehab increased for LEJR patients in the CJR and control groups, although more so for CJR patients (Exhibit 44). Our panel of clinical experts corroborated that industry trends supported the value of prehab for LEJR patients. Panelists, however, noted that the evidence is inconclusive as to the impact of prehab on post-surgical recovery time or functional outcomes, but they were unanimous that it is beneficial in providing education and "demystifying" the post-surgical rehabilitation process.

The goals of prehab differed across the CJR participant hospitals. Some site visit interviewees said that it was to strengthen deconditioned muscles around a patient's joint to improve surgical outcomes. Others said it was to educate patients about the exercises they would need to do after their surgery to facilitate quicker ambulation and reduce hospital LOS. Providers that did not promote prehab often indicated concerns about potentially exhausting patients' outpatient therapy benefit or that post-operative therapy was more important for patients' functional outcomes.

Exhibit 44: Greater increase in prehabilitation for CJR elective episodes than for control group episodes, PY1-2



Source: Lewin analysis of Medicare claims and enrollment data for episodes initiated on or after January 2012 that ended by December 2017.

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

Prehabilitation is defined as any outpatient physical therapy or occupational therapy visits during the 30 days before the anchor hospitalization.

The gray shading represents the 95% confidence interval for the CJR estimate.

PY = performance year.

b. Hospital stay

Flexibilities under the CJR model allow participant hospitals to share internal cost savings (ICS) with surgeons, which may spur focus on reducing internal hospital costs. While these cost reductions do not directly contribute to decreased payments under the model, sharing ICS can help hospitals engage or reward physicians for controlling episode payments.

From the Case Studies

Hospital G used gainsharing under the CJR model to motivate surgeons to participate in negotiation efforts with implant vendors. The hospital developed separate pricing schemes for high demand and low demand implants, which resulted in over $100,000 in savings to the hospital for implants through the second performance year of the model. The hospital interviewees described gainsharing agreements as "economic alignment."

Standardize implants

Reducing the costs of surgical implants often involved standardizing implants used by the hospitals' surgeons by, for example, contracting with only one or two implant vendors, implementing price ceilings or price caps for implants, negotiating contracts with high-volume vendors, or using group purchasing organizations (GPOs). Most interviewees noted that the key factor in successfully reducing implant costs was orthopedic surgeon support and engagement in these standardization activities. Several other interviewees noted that they had not yet standardized implants, but were negotiating with their surgeons to do so. Many noted that having better data on implant costs allowed them to set more accurate benchmark prices and to obtain surgeon buy-in regarding new implant procurement strategies.

The CJR hospital survey indicated that 59% of hospital respondents had implemented vendor practices to narrow their implant options. Half of respondents indicated the CJR model influenced their decision to implement that strategy. [35]

Reduce length of stay

Hospital interviewees described efforts to reduce LOS, such as changes to pain management and physical therapy (PT) services. The majority of interviewees began these efforts prior to the CJR model, which reflected physician preferences or best practices and were sometimes required by certification or accrediting organizations such as the Joint Commission.

[35] Combined survey responses of "A little", "somewhat", "very" or "extremely" influential are reported throughout the report.

> *"The doctors will tell you that pain management is what deserves credit for the reduction in length of stay. The hospital had about a 3.5 day length of stay and now it's down to 1.4 days. The new protocol has allowed patients to get up day-of-surgery and begin the healing process with therapy."*
>
> - Hospital Interviewee

Hospital interviewees often mentioned that standardizing pain management and changing intra-operative and post-operative pain management practices led to improvements in post-operative patient status. This, in turn, resulted in earlier and more intensive patient ambulation and, therefore, shorter inpatient LOS. Interviewees described working with anesthesia and surgical teams to standardize pain management practices. They also discussed engaging staff in pain management through surgical councils that reviewed system-level data and implemented best practices, and increasing collaboration with pain management services or case review groups.

Hospital interviewees frequently discussed ambulating patients on the day of surgery and multiple times per day, as well as implementing more aggressive PT plans than in the past. They said the hospital increased staff availability (e.g., on weekends, in evenings) and began orthopedic procedures earlier in the day to facilitate these changes. The majority of interviewees also mentioned that changes made to PT orders were applied to all patients, not just CJR patients.

The majority of hospital survey respondents (89%) implemented same day post-surgery ambulation and physical therapy for joint replacement patients and 92% implemented pain management practices that allow for early patient mobility. Roughly half of respondents indicated the CJR model was influential in their decision to implement both practices.

c. Post-discharge

Consistent with the findings from the PY1 Evaluation Report, hospitals reported focusing on discharging patients directly home (with home health or outpatient therapy) instead of to institutional PAC settings, improving overall quality, and preventing readmissions. To achieve these goals, hospitals extended patient follow-up for a longer period and developed PAC protocols and preferred PAC provider networks. Some of these activities were underway prior to the CJR model while the CJR model was a catalyst for others.

Patient follow-up

Hospital interviewees discussed efforts to strengthen and extend patient follow-up after discharge to reduce ED use and hospital readmissions. Commonly, interviewees mentioned calling patients within the first three days of hospital discharge and following up 30, 60, and 90 days post-discharge. While not as common, some interviewees said that they called patients weekly during the 90-day episode. Many mentioned having care coordinators systematically following up with SNFs on patient status or investing in data tracking and analysis software that allowed the hospital to follow patients. Common barriers to follow-up noted by interviewees included lack of dedicated care coordinators and difficulties in tracking patients once they were "back in their communities." Several interviewees described recently implemented or planned investments in infrastructure to help with patient follow-up after discharge.

From the Case Studies

At **Hospital F**, interviewees described the hospital's nurse practitioner care navigation program. Patients received the care navigator's phone number that they or their caregiver could call with any questions or concerns "24/7." The care navigator engages patients within 72 hours of discharge and follows patients after hospital discharge throughout the 90-day period. This follow up protocol was implemented to improve the hospital's ability to catch any signs of adverse events during the full episode of care.

At the time of the CJR hospital survey, 81% of respondents reported that follow up appointments were scheduled for all LEJR patients prior to discharge and 65% reported completing repeated telephonic follow-up during the entire 90-day episode. About half indicated the CJR model was influential in their decision to schedule follow-up appointments prior to discharge while 77% reported that the model was influential in telephone follow-up and tracking patients throughout the episode.

An In-Depth Look: Care Navigators

Many hospital interviewees and survey respondents reported staffing changes to accommodate CJR patients. Often, new staff were hired with titles such as care planner, case manager, navigator, or transition coordinator. Their responsibilities often included conducting an initial patient assessment two to four weeks before the surgery, which served as the "preliminary discharge plan," and their patient contact continued through the 90-day post-discharge period. Some of the care navigators followed the patients while they were in PAC facilities to prevent readmission and ensure timely and appropriate discharge. PAC staff viewed this arrangement favorably and appreciated the support with medication reconciliation and pain management. Many interviewees and survey respondents noted that having one consistent individual throughout the care pathway improved communication both within the team and between the surgeon and patient. One hospital survey respondent noted, "Assignment of a CJR patient navigator to follow the patients has been the most effective method of controlling triage to post-acute care and readmissions, through improved communication."

Data collection

While submitting patient-reported outcome (PRO) data is voluntary under the CJR model, hospitals can earn points towards a higher quality category of performance that will either increase reconciliation amounts or reduce CMS repayment amounts.[36] The collected data can also be used by participant hospitals for improving performance and reducing costs, thus increasing the performance on other CJR metrics, which could further increase payments.

> *"I think there are some things that are so high-yield and so fantastic about these bundles. They are pushing clinical excellence in ways that are so needed, and then there are these little tethers to it…and then if they [the patients] don't answer every question, it doesn't count. We could have been helping people with this time that we [spent collecting and uploading PRO data]…And there would have been better ways, I think, to get to that information than tethering it to this program."*
>
> - Affiliate health system representative

Consistent across the majority of interviews, hospital staff reported that collecting PRO data has been challenging. One hospital noted that PRO data collection was their "biggest stumbling block" and the data management team further revealed that PRO data "does not align with surgery process goals" and therefore feels disconnected to what the hospital is trying to accomplish. Reaching patients at the nine-month mark was noted as particularly problematic. Some interviewees reported that they mail paper surveys to their patients, but many of them have problems with comprehension and completion.

At the time of the CJR hospital survey, 50% of respondents reported that the hospital collects and reports PRO data in the electronic health record. Most respondents indicated that the CJR model influenced their decision to implement this (71%).

[36] To meet the PRO data submission requirements, hospitals need to submit the VR-12 or PROMIS Global-10 PRO measure as well as the full HOOS/KOOS or the HOOS JR / KOOS JR for patients undergoing eligible elective primary THA/TKA procedures. CJR hospitals need to submit PRO data for a minimum of 50% of the eligible cases or for 50 cases, whichever is most appropriate.

Preferred provider network

Many hospital interviewees indicated expending significant effort to identify the "highest quality" SNFs and improve the working relationship with those providers. The intent was to minimize SNF LOS and reduce readmissions through higher quality PAC. While many interviewees noted they had preferred provider lists prior to the CJR model, many described updates to their selection process under the model. The sophistication of the selection process for preferred providers varied from a "sense that they did well" to the use of algorithms and data metrics. Many hospitals reported relying on the CMS Five-Star Quality Rating Systems. Other hospitals selected SNFs with the shortest LOS. One hospital narrowed its preferred network through a Request for Information to local SNFs that asked them for metrics on responsiveness to referral requests, patient satisfaction results, LOS for joint replacement patients, and Medicare Star Ratings.

> *"And so now that we've been able to have that communication upfront with the patient and family, we have actually gone through and looked at what is each SNF's quality data and here are the ones that far outweigh the others. The patient, the family can still have their choice regardless, but it's a more well-informed decision for them."*
>
> - Hospital director of quality

Hospital interviewees often described requirements for participating preferred PAC providers. One of the more involved protocols required preferred PAC providers to attend weekly calls with the hospital, conduct root cause analysis of each hospital readmission, participate in quarterly evaluations, provide 24/7 coverage by registered nurses, and provide rehabilitation services every day of the week. Several hospitals required preferred SNFs to have software for read-only access to the hospital's electronic medical record (EMR).

Nearly two-thirds of CJR hospital survey respondents indicated that they had implemented a preferred provider network for PAC providers (62%) and 72% indicated the CJR model influenced their decision to implement or enhance this strategy.

Hospitals often reported that honoring patient choice made it difficult to guide patients to preferred providers because patients often made the selection based on convenience (e.g., proximity to their home) or a recommendation of a friend or family member. To influence their choice, hospital staff informed patients about their PAC options and emphasized the preferred PAC providers during pre-surgical educational classes, direct engagement with the navigators, or putting the preferred providers at the top of PAC lists along with their Star Rating. In other instances, surgeons encouraged patients to consider particular SNFs. One hospital reported that the CJR model motivated them to collect information about why patients chose non-preferred providers so that they could better influence the choice of PAC provider.

An In-Depth Look: Did hospital actions result in reduced episode payments?

The hospital survey conducted in PY2 asked CJR participant hospitals if they were implementing or planned to implement 14 different care redesign activities, which included post-discharge telephonic follow-up with patients and creation of preferred PAC provider networks. The survey also asked hospitals if the CJR model had influenced their implementation or enhancement of these activities. We summarized the questions into two indices – the care redesign implementation index and the influence of the CJR model on care redesign index.

We found that respondents reported implementing *nearly two-thirds* of the care redesign activities and that the CJR model was influential in their decision to implement care redesign. Hospitals reported that on average *60% of the care redesign activities* they currently had implemented or planned to implement were *influenced in some way* by the CJR model.

Additionally, we explored the relationship between these indices and two measures of hospital financial performance under the CJR model: average change in episode payments from baseline to PY2 and receiving reconciliation payments. We found a significant correlation between change in episode payments and the number of care redesign activities implemented by hospitals (correlation coefficient, -0.24; $p<0.01$). In other words, the greater the number of care redesign activities implemented by the hospital, the greater the reduction in payments. With respect to individual activities, hospitals that reported implementing classes prior to the joint replacement admission (correlation coefficient, -0.26; $p<0.01$) and regular reporting of patient outcomes to surgeons (-0.23; $p<0.01$) experienced greater reductions in payments.

Further, we found a significant correlation between change in episode payments and the influence of the CJR model on care redesign implementation (correlation coefficient, -0.26; $p<0.01$). Hospitals that reported any influence of the CJR model on their implementation of repeated telephonic follow-up and tracking (-0.25; $p<0.01$), preferred PAC provider networks 0.26; $p<0.01$), and regular meetings between hospital and PAC providers (-0.23; $p<0.01$) experienced greater reductions in payments.

d. Conclusion

CJR hospitals reported a range of enhanced or new initiatives implemented as part of the pre-surgical, inpatient, and post-discharge care pathways for LEJR patients. Participating hospitals reported an increased focus on educating patients, providing physical therapy earlier and more often, using data to inform clinical decision-making, and working with surgeons and PAC provider partners to adopt more efficient practices. These efforts were undertaken to shift PAC to less expensive settings and reduce institutional PAC lengths of stay.

K. How were relationships with orthopedic surgeons and PAC providers impacted by the CJR model?

Because the CJR model holds hospitals accountable for an LEJR episode of care that extends for 90 days post-discharge, participant hospitals may establish or strengthen relationships with other health care providers involved in the episode. Orthopedic surgeons and PAC providers, in particular, may influence the cost and quality of care across the LEJR episode. Whether and how

hospitals work with orthopedic surgeons and PAC providers may affect their success in reducing payments and improving quality under the model. We examined how these relationships have changed due to the CJR model and the factors that affect these relationships.

1. Key findings

Representatives from hospitals reported:

- Sharing CJR performance data and gainsharing helped engage physicians in their hospital's activities in response to the model
- Having a designated physician champion to participate in CJR model meetings, review quality data, and serve as a liaison between the hospital and other surgeons about performance outcomes

Representatives from PAC providers reported:

- Increased collaboration with other health care providers, including hospitals, orthopedic surgeons, primary care providers, and other PAC providers
- Changing care pathways and protocols for LEJR patients to help prevent hospital readmissions in response to the CJR model

2. Methods

We synthesized information from site visit interviews, telephone interviews, and the hospital survey to understand the changes in relationships between hospitals, orthopedic surgeons, and PAC providers in response to the CJR model. More detail on these data sources is available in Appendix E.

3. Results

a. Relationships with orthopedic surgeons

Physician involvement in care redesign

Hospital interviewees described engaging orthopedic surgeons in efforts to redirect patient discharge destination from SNFs to home and improve care coordination after discharge to reduce readmissions. These efforts included discussing discharge destination with patients during pre-operative office visits, starting the discharge planning process earlier in the care pathway, encouraging patients to attend pre-surgical education classes, implementing interdisciplinary rounding involving surgeons and physical therapists, and improving coordination between the emergency room physicians and surgeons. Many hospital interviewees also described working with surgeons to standardize implants, order sets, or clinical pathways.

Strategies for increasing engagement

> *"It is a very transparent dashboard where the surgeons can see how they compare to other surgeons. It gives [results] by DRG – the total number of patients in that DRG by physician…the number of patients, the average length of stay, and the percent of where those patients' discharge disposition was. It also gives the number of readmissions for surgeon for the 30 and 90 day periods."*
>
> - Hospital Interviewee

Interviewees reported that sharing CJR performance data was a critical strategy for engaging physicians in their hospital's activities related to the model. Nearly all of the hospitals interviewed reported sharing information about episode costs, quality measures, and PAC utilization. A few shared information about patient outcomes, such as readmissions or length of stay, on a monthly or quarterly basis and less frequently, total episode costs. Some interviewees reported sharing only aggregate data with surgeons, but others reported sharing physician-level data, such as cost per case, average length of stay, readmission rates, and discharge destination in an "un-blinded" manner with all surgeons, even when surgeons were in different group practices.

Interviewees said that the CJR performance data was particularly effective in helping them work with surgeons to shift discharge destinations and mentioned that surgeons were more likely to discharge patients to lower intensity settings after viewing episode data. They indicated that discharge patterns and episode cost data were effective in demonstrating the value of ordering outpatient rehabilitation versus SNF or HHA care. As one interviewee reported, "when they saw the actual difference in costs, it shifted a lot of practice."

The majority (73%) of hospitals that responded to our survey indicated that they were reporting patient outcomes to individual surgeons, and most of those hospitals (77%) indicated that the CJR model influenced their decision to implement the strategy. Over half (61%) of respondents felt that physician engagement in care redesign activities had improved since the CJR model was implemented.

Gainsharing

About half of the hospital representatives we interviewed by telephone stated that they were planning to implement gainsharing agreements with orthopedic surgeons in the future or currently had gainsharing agreements in place; the other half reported that they had no plans to gainshare with surgeons.[37] Hospitals in MSAs with a large supply of orthopedic surgeons were more likely to report that they planned to

From the Case Studies

Hospital J's service line administrators explained their experience implementing gainsharing agreements with orthopedic surgeons at their hospital. They described that the highly competitive labor market influenced their decision to begin gainsharing, as the surgeons are highly respected and service line leadership felt it was necessary to keep them happy and continue practicing at Hospital J.

[37] The fourth round of telephone interviews was conducted with 46 hospitals in the spring of 2018. It focused on hospitals' relationships with orthopedic surgeons and how those relationships evolved since the start of the CJR model.

gainshare. Among those with gainsharing agreements in place, most indicated that gainsharing was their most successful strategy to engage physicians in their hospital's response to the CJR model. Many reported that volume of LEJR surgeries performed at the hospital influenced their decision to enter into gainsharing agreements.

> *"If you don't have 100+ cases a year we'd advise against doing gainsharing. We need to make sure we have the beneficiary notification in place, the website information in place, the collection of the data and the calculations to write the check. It's a pretty burdensome process, if you aren't doing enough cases it probably isn't worthwhile."*
>
> - Health System Administrator

When asked about details of their gainsharing agreements with surgeons, interviewees most commonly reported that they shared both reconciliation payments and internal cost savings, but none reported sharing repayments. Interviewees described specific quality and utilization thresholds, like targets for complication rates, 90-day readmission rates, or use of SNF, which surgeons had to meet to share in savings. Gainsharing agreements at a few hospitals included requirements that surgeons participate in activities related to the hospital's response to the model, such as patient attendance at pre-operative classes, or surgeons' compliance with the hospital's preferred implant list.

Most interviewees described some challenges in designing and implementing gainsharing agreements. Common challenges included reaching consensus on contract language with individual surgeons and leadership of surgery groups, administrative burden, and selecting appropriate and reliable performance metrics.

A few interviewees noted that gainsharing agreements did not provide financial incentives sufficient to motivate surgeon behavior. Instead, they felt that surgeons responded positively to the formal expression of partnership that was established through the agreements, resulting in increased support for and participation in care redesign activities. Overall, interviewees conveyed the perception that gainsharing increased physician engagement, willingness to implement new clinical care processes, and collaboration to achieve lower total episode costs.

> *"He's the voice. When this all began he participated in all of the interdisciplinary meetings so we could get his feedback and he could support us. He shares the information with his peers as well, just to be a physician speaking instead of nursing"*
>
> - Hospital Interviewee

Physician champion

The majority of the hospitals we spoke with reported that they have a designated physician champion for their hospital's response to the CJR model. The level of involvement in CJR model activities varied but hospital interviewees most commonly identified the champion's responsibilities as attendance at CJR model meetings (and sometimes leading them), reviewing quality data, and serving as a liaison between the hospital and other surgeons about performance outcomes.

b. Relationships with PAC providers

During site visits, the evaluation team conducted interviews with SNF and HHA representatives to better understand their perspective on how the CJR model affected hospital relationships. While themes regarding key changes were identified, often it was not possible to determine the specific influence of the CJR model on the activities discussed due to interviewee's limited tenure in position, limited perspective (i.e. outside of scope of interviewee's role), or inability to distinguish the influence of concurrent initiatives. Future evaluation efforts will aim to better understand the specific influence of the CJR model as well as the resulting outcomes of key changes.

Communication and coordination

Similar to findings from hospital interviews conducted during site visits, PAC provider interviewees reported increased collaboration with other health care providers, including hospitals, orthopedic surgeons, primary care providers (PCPs), and other PAC providers regarding LEJR patients. Interviewees reported greater communication between the PAC staff and a point person at the hospital, such as a transition coordinator, care manager, or social worker. Many saw value in meeting with hospitals on a regular basis to share and discuss data and discharge plans, review readmissions, or raise any concerns about patients' progress.

Increased coordination with orthopedic surgeons and PAC providers as a result of the CJR model was another common theme. Interviewees described having the ability to call the orthopedic surgeon directly to consult on care plans, if needed. A few interviewees also reported that the orthopedic surgeon comes to the PAC setting to visit patients. PAC provider interviewees also discussed communicating and coordinating with patients' PCPs. For instance, a few interviewees reported that they send discharge summaries to the patients' PCPs on their patients' status. Others described having staff available, typically a social worker, to help schedule patients' follow up appointments with their PCPs.

At the time of the CJR hospital survey, 45% of hospitals reported implementing regular meetings with PAC providers to share financial or clinical status updates for LEJR patients and 73% felt that the CJR model was influential in the decision to implement these changes. Sixty-eight percent of hospital respondents reported that their communication with PAC staff improved since the beginning of the model.

Only 27% of hospital survey respondents allowed PAC providers to access LEJR patients' EMRs. Some PAC provider interviewees discussed the ability to receive certain documents from hospitals' EMRs that could be scanned into their system while others mentioned duplicating efforts because of needing to manually enter the patient's information into their own systems. This lack of interoperability across EMRs was noted as a challenge.

Care pathways

> *"So what we did on our end to really meet the needs of the patients and kind of meet the needs of the CJR (…) is we hired a physical therapist that works from 1:00 pm to 9:00 pm. Just so when they do come into the building we're doing day one evaluations. We have a true seven day per week program where I do have PTs and PTAs working both Saturday and Sunday."*
>
> - SNF Interviewee

The majority of PAC provider interviewees reported some changes in their care pathways for LEJR patients. About half said these changes were due to the CJR model, while others reported these changes were driven by specific requests from hospitals or physicians. Some PAC providers described offering therapy seven days a week and increasing the number of days, duration, or sessions each patient received. Among HHA interviewees, over half reported that they frontload therapy sessions in the first two weeks for their patients; however, most did not specify whether this was a change due to the CJR model.

PAC provider interviewees also reported changing their protocols to help prevent hospital readmissions in response to the CJR model. Some interviewees described steps that they take while the patient is still in the PAC setting, such as frequently reviewing potential patient complications and consulting with providers in the hospital via phone or telehealth monitoring systems. Other PAC provider interviewees described implementing post-discharge protocols, such as conducting follow-up phone calls with patients and offering direct admission back into the PAC setting from the community.

An In-Depth Look: Patient Messaging

SNF interviewees described changing messaging strategies to LEJR patients and the important role it played in the effort to shift the patient mindset around length of stay. Nearly all SNF interviewees described consistently communicating expectations to patients around length of stay, and the importance of initiating this conversation prior to surgery and maintaining it across the care pathway. Interviewees also noted that consistent and frequent messaging from surgeons, direct care staff, social workers, and case managers was critical to gain patient buy-in and reduce length of stay.

c. Conclusion

There are indications that CJR participant hospitals engaged with orthopedic surgeons and PAC providers to change the services provided to LEJR patients. Enhanced communication and information sharing focused on reduced hospital length of stay and readmissions and minimizing institutional PAC use. Interviewees were not always able to indicate the extent to which these activities were due to the CJR model or whether they were effective. Future evaluation efforts will build on this these findings and focus on understanding what impacts resulted from changes.

III. Discussion and Conclusion

A. Discussion

The CJR model tests whether episode-based bundled payments and quality measurement for LEJR can lower payments and improve quality through improved care coordination across all providers involved in the episode. The evaluation of the first two performance years indicates that the CJR model was successful in reducing average episode payments. Quality of care was maintained, even with the significant decreases in institutional PAC use that led to the reduction in payments. Furthermore, after accounting for reconciliation payments paid to participant hospitals, the CJR model likely resulted in savings to the Medicare program, although there is a wide range around the estimated savings.

Because of its design as a mandatory, randomized model, this evaluation was able to assess the impact of episode-based payments across a variety of participant hospitals. Average episode payments decreased in MSAs with historically high and historically low payments, for hospitals with high and low LEJR episode volume, for patients with planned LEJRs and those with LEJR due to fracture, and for patients with varying levels of complexity. This suggests that lower episode payments can be achieved across a variety of LEJR episodes and that hospitals in a range of circumstances can meet the goals of the CJR model.

Hospital representatives and clinicians considered the entire episode of care in response to the CJR model incentives. They indicated that they focused on reducing institutional PAC by emphasizing patient education and earlier discharge planning, speeding physical therapy after the surgery, and working with PAC providers on care protocols. Hospital representatives also indicated increased collaboration with physicians and PAC providers to reduce spending and improve quality of care.

Changes in PAC use, which suggest shifts to less intensive sites of care, contributed the most to the decrease in episode payments. Discharges to IRF went down. The proportion of LEJR patients discharged home with home health care rose. Although the proportion of LEJR patients first discharged to SNF remained relatively constant, the average length of SNF stays decreased, which resulted in lower payments. Evidence suggests that CJR participant hospitals may have better targeted where their patients were discharged, reserving institutional PAC for patients with greater needs. Shifts in PAC use did not seem to adversely affect longer-term functional status and pain outcomes; CJR and control group survey respondents reported similar gains in functional status from before their hospitalization to after the end of the episode. Changes in functional status and pain outcomes during the shorter term, however, indicated a reduction in the proportion of CJR patients discharged to PAC who improved their functional status during their PAC stay relative to control patients. For CJR patients discharged to SNF, this was in part due to shorter SNF stays. Orthopedic surgeons and other clinicians we interviewed as well as clinical review panel members were consistent in their view that home was the best place for most patients to recover and that with time, CJR patients would achieve the same level of functioning after their surgery.

There are no consistent indications that the CJR model resulted in participant hospitals lowering their average episode payments by changing their mix of patients to those likely to require fewer services, refusing care to the most complex patients, or increasing the volume of episodes. The CJR model incorporates features to guard against these potential unintended consequences, in particular the four separate target prices and the MSA-wide implementation. Some changes in patient characteristics for those with elective episodes without major complications or comorbidities, however, may indicate a less severe patient mix, which warrants continued scrutiny.

Hospital representatives that we interviewed often expressed positive outcomes from participation in the mandatory model. They noted reductions in internal costs, although these cost reductions do not directly contribute to decreased payments under the model. This indicates, however, that the CJR model may have provided renewed focus on improving internal efficiencies, possibly through better data or the ability to better engage physicians, potentially through gainsharing.

Lower episode payments under the model result in savings to the Medicare program only if the aggregate reduction was greater than the reconciliation payments paid to participant hospitals. After accounting for the reconciliation payments, as well as the confidence interval around the estimated decrease in episode payments and the volume of CJR episodes, the CJR model resulted in estimated savings of $17.4 million, however the estimate ranges from losses of up to $41.2 million to savings of up to $75.9 million. Calculated another way, we are 69% confident that Medicare achieved savings under the CJR model. The probability of savings would have increased to 79% had downside risk not been waived during the first performance year of the model. While this may lead some to conclude that Medicare could lower quality-adjusted target prices to ensure Medicare savings, we do not know if participant hospitals would have made similar decisions about how to respond or reduced episode payments under different model design specifications.

B. Considerations

The CJR model's mandatory, randomized design reduces self-selection bias and enhances generalizability, while implementation at the MSA level constrains opportunities for patient selection. This design mitigates some of the most important factors that have hampered the evaluation of previous, voluntary episode-based bundled payment models.[38] However, there are remaining data limitations and unobserved factors that may temper the strength of our conclusions. Furthermore, it is essential to be aware of what can and cannot be concluded from a given result as our level of confidence and ability to generalize across results varies.

To address these concerns, we have employed a robust mixed methods approach that assesses the impact of the CJR model through multiple types of analyses. This approach allows results to be triangulated across data sources and methods, with shortcomings or open questions from one

[38] Gronniger T, Fiedler M, Patel K, Adler L, Ginsberg P. How should the Trump Administration handle Medicare's new bundled payment programs? Health Affairs blog. April 2017.

analysis addressed by another. Taken together, the results presented in this report point to the promise of episode-based payment models, subject to the following considerations and caveats.

Our evaluation includes numerous outcomes, which increases the risk that some of our statistically significant findings are due to chance. While we plan to implement adjustments to account for this potential statistical problem, the strong statistical significance of many of our results means they are unlikely to be affected by this issue. In addition, the certainty of our conclusions is bolstered because results tell a consistent story across the various methods we have employed.

The analysis of the site visit and telephone interview data provide descriptions of common themes that emerged. For both site visits and interviews, we oversampled hospitals that had high average episode payments relative to their quality-adjusted target price. This was to ensure that we captured information relative to the widest range of strategies implemented in response to the model from hospitals with the most pressure to respond and as such limits generalizability to the broader population.

The estimates in this report may also be affected by recent changes in the model. The proposed rule for the policy that modified the CJR model and made it voluntary in 33 MSAs starting Jan 1, 2018 was released on Aug 17, 2017. Some hospitals may have stopped responding to the model's incentives if they knew they were not going to opt-in. This effect would most likely bias our payment impact estimates towards zero, leading to an underestimate of the CJR model's financial impact. At this time, at most 6.9% of all CJR intervention episodes could have been affected. In future analyses we will explore the possibility of controlling for and estimating the magnitude of this effect.

Taken together, quantitative results from claims, patient assessments, and patient surveys combined with information gleaned from site visits, provider telephone interviews, and hospital surveys provide a strong evaluation of the CJR model. Consistency across findings lends strength to our conclusions, while inconsistencies raise questions for further inquiry.

C. Conclusion

This second annual evaluation report demonstrates that the CJR model continues to be a promising approach to reducing payments for an episode of care that begins with LEJR surgery. This evaluation indicates that a range of hospitals, with a range of resources and circumstances, can and do respond to the incentives under a mandatory episode-based payment approach for LEJR episodes to reduce per episode payments for both planned LEJR episodes and those due to fracture. In response to the CJR model, participant hospitals said they continued with care redesign efforts and engaged in strategies to change PAC use after hospital discharge. Even with lower service use under the CJR model, there were few indications that quality of care changed, which suggests that hospitals focused on reducing services that were of marginal or no clinical value.

In future reports we will expand our understanding of the payment decreases under the CJR model and whether the lower episode payments translate into savings for the Medicare program. With

additional time under the model, we will have more information to evaluate whether relative differences in functional improvement persist and clinical opinions on the importance of any differences. As the quality-adjusted target prices shift to be based more on regional historical averages, it will be particularly important to evaluate relative performance of participant hospitals that have historical payments above their peers as well as those with historical payments below their peers.

This annual report provides results for the first two performance years of the CJR model when participation was mandatory in all 67 MSAs. The design of the model changed in PY3 with hospitals in the 34 MSAs with the highest average historical episode payments remaining mandatory. Hospitals in the other 33 MSAs and rural and low-volume hospitals were no longer required to participate, although they could choose to opt in to the CJR model for the final three performance years. This change in the model design offers unique opportunities to add research questions about differences in hospitals that opted to stay in the model versus those that took the opportunity to leave. Another important area of inquiry this opens up is whether changes made in response to the model persist when the hospital will no longer be able to earn reconciliation payments because they are no longer in the CJR model. Further, we will be able to examine additional differences in market level effects of the model by comparing MSAs that remain fully mandatory with those that have a mixture of continuing versus exiting hospitals.