CMS Comprehensive Care for Joint Replacement Model: Performance Year 3 Evaluation Report

Third Annual Report

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

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

November 2020

CMS Comprehensive Care for Joint Replacement Model: Performance Year 3 Evaluation Report

Third Annual Report

Table of Contents

List of Appendices

Executive Summary

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

The third annual CJR model evaluation report presents findings from the first three performance years, which include episodes initiated on or after April 1, 2016 that ended by December 31, 2018. At the start of performance year (PY) 3, the number of mandatory Metropolitan Statistical Areas (MSAs) was scaled back from the 67 original randomly selected MSAs to the 34 MSAs with the highest average historical payments.[3] This report focuses on the 378 mandatory CJR hospitals in the 34 mandatory MSAs that were continuously required to participate through the entire model. This report does not assess the impact of the CJR model on hospitals designated as low-volume or rural in mandatory MSAs or on hospitals in the 33 voluntary MSAs, even though they participated for some or all of the first three performance years. These hospitals will be evaluated in future annual reports.

In the first three performance years, mandatory CJR hospitals achieved a statistically significant reduction in average episode payments due to reductions in institutional post-acute care (PAC) use. After accounting for reconciliation payments made to mandatory CJR hospitals, the payment reductions made by mandatory CJR hospitals resulted in Medicare savings during the first three performance years, although this does not consider the experience across all hospitals that ever participated in the CJR model. Quality of care, as measured by the unplanned readmission rate, emergency department use, and mortality, improved or was maintained under the CJR model. Further, 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 market conditions like availability of quality post-acute care providers and financial relationships with orthopedic surgeons affected their ability to respond to the model.

---

[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. Hospitals are excluded if they are currently participating in a Bundled Payments for Care Improvement model for LEJR or are cancer hospitals.

They said that interdisciplinary teams coordinated care for LEJR patients throughout the episode, although these activities were not necessarily in response to the CJR model.

A. Introduction

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

The CJR model originally required hospitals in 67 MSAs to participate. Because of the CJR model's mandatory and randomized design, a spectrum of hospitals with varying levels of infrastructure, care redesign experience, episode costs, utilization, and market positions participated, which allowed a broad test of the CJR model. In the third performance year, beginning January 2018, CMS scaled back the number of mandatory MSAs to the 34 MSAs with the highest average historical episode payments. Hospitals in these mandatory MSAs that were not designated as low-volume or rural (mandatory CJR hospitals) were required to continue their participation in the CJR model.[4] CMS provided a one-time opportunity for hospitals in the 33 MSAs with lower average historical payments (voluntary MSAs) and all hospitals that were designated as low-volume or rural a one-time opportunity to opt-in to the CJR model for PY3-5.

This evaluation report focuses on the 378 mandatory CJR hospitals; it does not assess the impact of the CJR model on the hospitals that were given the opportunity to voluntarily participate. Of the hospitals in the voluntary MSAs, 76 opted to continue participation in the CJR model and 233 did not opt-in for performance year 3. Hospitals that opted to continue participating likely did so because they were well positioned to receive reconciliation payments under the CJR model. This report does not analyze hospitals that were given the option to continue participating to focus on the evaluation of a mandatory model.

This evaluation accounts for the effect of changes to Medicare coverage of LEJR procedures that affected all hospitals starting in performance year 3. In 2018, Medicare began covering TKAs performed in the hospital outpatient department, so Medicare pays for TKAs performed in the hospital inpatient (IP) or outpatient (OP) setting. When a TKA is performed on an inpatient, the hospital's payment is made under the Inpatient Prospective Payment System (IPPS) and the surgery triggers a CJR episode for CJR participating hospitals. When performed on an outpatient,

---

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

the hospital's payment is made under the Outpatient Prospective Payment System (OPPS) and the surgery does not trigger a CJR episode. In this report, we explore whether the response to the new coverage policy differed between mandatory CJR and other hospitals and how any differences affect the CJR model impact analyses presented in this report.

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

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

B. Results

The CJR model resulted in decreases in average payments (standardized allowed amounts) for LEJR episodes at mandatory CJR hospitals during the first three performance years. The amount of the reduction, however, was affected by the change in Medicare coverage of outpatient TKAs, because mandatory CJR hospitals responded to this policy change differently than control group hospitals. Mandatory CJR hospitals shifted a smaller portion of TKA surgeries to the outpatient setting than control group hospitals (about 19% vs. 29%). To account for the differential response to the coverage of TKAs in the hospital outpatient setting, we estimated the impact of the CJR model with two samples: one that included all LEJRs, including inpatient and outpatient TKAs even though outpatient TKAs do not result in episodes under the CJR model; and one that included only LEJRs performed in the inpatient setting, which reflect actual episodes under the CJR model. The estimates based on all LEJRs (IP+OP) underestimate the impact of the CJR model, while the impact estimates based on inpatient LEJRs overestimate the impact of the CJR model. Because our analyses suggest that the lower outpatient TKA rate in mandatory CJR hospitals was due to the CJR model and not other factors, the estimate based on all LEJRs (IP+OP) is closer to the truth because it considers the differential response to the outpatient TKA coverage policy. For all LEJRs

> *Baseline* refers to a three-year period (2012 through 2014) prior to the CJR model that serves as a benchmark against which performance under the CJR model is compared.
>
> *Intervention* refers to the period that the CJR model was in effect and studied through our analyses. Except where noted, the intervention period includes the first three performance years of the CJR model (April 2016 through December 2018).

(IP+OP), payments decreased by $1,378 more for mandatory CJR hospitals than for control group hospitals, or 4.7% from CJR baseline payments, although this underestimates the impact of the model (p<0.01).[5] Based on the other sample of inpatient LEJR episodes, payments decreased by $1,540 more for mandatory CJR hospitals than for control group hospitals, or 5.3% from CJR baseline payments (p<0.01), although this is an overestimate of the impact of the model.

The statistically significant decreases in payments in both episode samples were primarily due to reductions in institutional PAC use. Mandatory CJR hospitals discharged a relatively smaller proportion of patients to an inpatient rehabilitation facility (IRF) or a skilled nursing facility (SNF), and a relatively larger proportion of patients to a home health agency (HHA) than control group hospitals. Furthermore, CJR patients with a SNF stay spent relatively fewer days in a SNF than control group patients. These changes in utilization resulted in statistically significant decreases in IRF and SNF payments, which drove the decrease in average episode payments.

The CJR model resulted in Medicare savings for mandatory CJR hospitals during the first three performance years. Medicare savings is the difference between the reduction in episode payments and reconciliation payments to and repayments from mandatory CJR hospitals. Net reconciliation payments were $90.5 million during the first three performance years.[6] Medicare savings estimates are presented as a range because the estimates are based on our statistical analysis of the reduction in payments that includes an upper and lower estimate. For all LEJRs (IP+OP), Medicare savings for mandatory CJR hospitals during the first three performance years was $61.6 million, which is an underestimate, that ranges from savings of $2.3 million to $120.9 million. For inpatient LEJR episodes, Medicare savings was $76.3 million, which is an overestimate, that ranges from savings of $16.0 million to $136.7 million. As explained earlier, the underestimate based on all LEJRs (IP+OP) is closer to the actual amount.

Quality of care, as measured by the unplanned readmission rate, emergency department use, and mortality, improved or was maintained for mandatory CJR hospitals. For patients who were first discharged to an IRF or HHA following their hospital stay, improvement in functional status and pain during their PAC stay were generally similar for CJR and control group patients, but changes were mixed for CJR patients who were first discharged to a SNF. By the end of the episode, CJR and control patient survey respondents reported similar functional status gains and pain levels from before the episode to after the end of the episode. CJR and control survey respondents reported similar satisfaction with overall recovery and care management, although CJR respondents reported worse care transition experiences and required more caregiver help than control

---

[5]   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.

[6]   The $90.5 million is based on $109.4 million in reconciliation payments made to mandatory CJR hospitals less $18.9 million in repayments received from mandatory CJR hospitals.

respondents. Orthopedic surgeons, other clinicians, and hospital staff we interviewed or consulted recognized that shifting recovery to the home setting may increase caregiver responsibilities, and were consistent in their view that home was the best place for most patients to recover, congruous with the research base.[7,8]

The evaluation also examined whether the CJR model resulted in any unintended consequences. For example, CJR participant hospitals could increase the volume of LEJRs to maximize reconciliation payments, delay services until after the end of the episode to keep episode payments low, or favor less complex patients who may be less costly. The CJR model was designed to protect against these responses by including all hospitals in the MSA, using a long episode definition, and other means, but the CJR model may still result in unintended consequences. We found no evidence that the CJR model was associated with an increase in the total market volume of LEJRs (inpatient plus outpatient). For inpatient LEJRs, patient complexity increased for patients receiving elective LEJR who did not have complications or comorbidities for both mandatory CJR and control group hospitals, although mandatory CJR hospitals experienced a smaller increase. As a result, relative to the control group, the population of CJR patients receiving elective LEJR without major complications or comorbidities was healthier in the intervention period than the baseline period.

CJR participant hospital representatives we interviewed indicated that hospital resources – such as relationships with surgeons, PAC providers, and health systems – and market conditions affected responses to the CJR model. In particular, hospitals were able to exert more control over care redesign when orthopedic surgeons were employed by or had a financial relationship with the hospital. Interviewees also noted that care coordination was more challenging when there was a low supply of quality PAC providers in the market. Despite variation in circumstances, hospital interviewees identified common approaches to care coordination for LEJR patients, although these activities were not consistently attributed to the CJR model. Interviewees indicated they began discharge planning earlier under the CJR model, often noting it occurred in conjunction with pre-surgical patient education. Interviewees also said they coordinated care for patients across the entire episode, tailoring approaches based on patient needs, as appropriate. At the same time, hospital and PAC provider interviewees noted that reducing the use of PAC could cause challenges for some patients or caregivers.

These key findings reflect our evaluation framework, which focuses on the hospital where the LEJR episode begins and recognizes the influence of financial, hospital, and market factors on hospitals' responses to the CJR model. Accordingly, the research questions considered under this evaluation were organized into five categories: (1) study population and policy context, (2) impact

---

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

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

of the CJR model, (3) financial risk or opportunity, (4) resources and market conditions, and (5) CJR participant hospital responses. Additional details about key findings are summarized under the main research questions addressed in this report.

1. Study population and policy context

a. How did the characteristics of CJR participant hospitals compare with those of other hospitals?

■ **Mandatory CJR hospitals had a wide range of characteristics similar to all other IPPS hospitals.** Mandatory CJR hospitals were distributed across the U.S. and they varied in size, LEJR volume, and patient complexity. The variation in characteristics for mandatory CJR hospitals was similar to that for all other IPPS hospitals, indicating that a wide spectrum of IPPS hospitals were represented by mandatory CJR hospitals. Compared to all other IPPS hospitals in the U.S., mandatory CJR hospitals were more likely to be larger and teaching hospitals, and located in the Northeast and South. Historically, they had higher average episode payments, more complex patient populations, and discharged a greater proportion of patients to institutional PAC settings. These differences were likely driven by the design of the model, which was rolled out in metropolitan areas with large populations and only hospitals in the 34 higher payment MSAs were required to continue participation in the CJR model in PY3.

■ **Hospitals in voluntary MSAs that opted-in to the CJR model likely did so because they anticipated receiving reconciliation payments**. In voluntary MSAs, opt-in hospitals were more likely to have started the CJR model with average historical episode payments below their quality-adjusted target price than non-opt-in hospitals, which means they would not have to lower payments to receive reconciliation payments. Opt-in hospitals were more likely to earn reconciliation payments in the first two years of the model than non-opt in hospitals.

b. Did CJR participant hospitals respond differently than other hospitals to Medicare coverage of TKA in the hospital outpatient setting and how would this affect the evaluation of the model?

■ **Mandatory CJR hospitals shifted a smaller proportion of TKAs to the outpatient setting than control group hospitals, which could result in lower average episode payments in the mandatory CJR hospitals relative to control group hospitals.** In 2018, mandatory CJR hospitals performed approximately 19% of their TKAs in the outpatient setting, while control group hospitals performed about 29% of their TKAs in the outpatient setting (p<0.01). Evidence suggests that this difference cannot be explained by other hospital or patient factors, so is likely due to the CJR model. That is, some inpatient TKA surgeries, which triggered CJR episodes, likely would have been outpatient procedures, which would not have triggered CJR episodes, absent the CJR model. Consequently, impact estimates based only on inpatient LEJRs are biased and too high for two reasons. First, CJR TKA episodes that likely would have been outpatient absent the model are incorrectly compared with inpatient control group TKAs, which

have higher payments, instead of outpatient control group TKAs, which have lower payments. Thus, the impact estimates would not account for the higher inpatient payments associated with LEJRs that would have been outpatient LEJRs absent the CJR model. Second, the differential response to the change in TKA coverage may have contributed to a relative decrease in the complexity of the average CJR patient, which our impact estimates may not fully account for despite our rigorous risk adjustment methodology. Because lower complexity patients are likely to incur lower payments, a relative decrease in CJR patient complexity due to the differential response to the outpatient TKA coverage policy would result in an overestimate of the impact of the model. Therefore, in this annual report we provide impact estimates based on all LEJRs, which include outpatient TKAs that were not episodes under the CJR model. Impact estimates based on all LEJRs (IP+OP) likely underestimate the impact of the CJR model. The two sets of estimates provide an upper and lower bound on the impact of the CJR model, although the estimate based on all LEJRs (IP+OP) is closer to the truth because it considers the differential outpatient TKA rate.

2. *Impact of the model*

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

■ **For mandatory CJR hospitals, there were statistically significant reductions in average episode payments during the first three performance years of the CJR model.** [9] Average payments for *all LEJRs (IP+OP)* decreased by $1,378 more than for the control group. This underestimated relative reduction in payments equates to a 4.7% decrease in average episode payments for *all LEJRs* from the baseline (p<0.01).[10] Average payments for *inpatient LEJR episodes* decreased by $1,540 more for CJR than for control group episodes. This overestimated relative reduction equates to a 5.3% decrease from the baseline.

■ **For hospitals in the original 67 CJR MSAs, there were statistically significant reductions in average episode payments during the first three performance years of the CJR model.** Average payments for *all LEJRs (IP+OP)* decreased by $866 more than for the control group. This underestimated relative reduction equates to a 3.2% decrease from the CJR baseline (p<0.01). Average payments for *inpatient LEJR episodes* decreased by $991 more for CJR than for control group episodes. This overestimated relative reduction equates to a 3.7% decrease from the baseline.

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

■ **Across the first three performance years, Medicare realized savings from mandatory CJR hospitals. Medicare likely did not realize savings in PY3, however,**

---

[9] 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.

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

due to the differential response to Medicare coverage of OP TKA. For the first three performance years, estimated Medicare savings from mandatory CJR hospitals was $61.6 million, ranging from savings of $2.3 million to $120.9 million. This Medicare savings estimate is based on an estimated reduction in total non-standardized paid amounts for *all LEJRs (IP+OP)* of $152.1 million (ranging from $92.7 million to $211.4 million) less $90.5 million in reconciliation payments made to or received from mandatory CJR hospitals.[11],[12] The estimates are equivalent to a savings of $536 per episode, ranging from savings of $20 to $1,052 per episode. This equates to savings of 2.0% from the baseline. While the estimate based on *all LEJRs (IP+OP)* likely underestimates Medicare savings, as discussed earlier, it is more representative of the true impact of the model than an estimate based on only IP LEJR episodes. The overestimate of Medicare savings based on *inpatient LEJR episodes* was similar in magnitude. For the first three performance years, estimated Medicare savings based on inpatient LEJR episodes was $76.3 million, ranging from savings of $16.0 million to $136.7 million. The *inpatient only* savings estimate is equivalent to $685 per episode, ranging from a savings of $143 to $1,227 per episode. This equates to savings of 2.5% from the baseline. While estimated savings for the first three years was similar between the two samples, the estimates differed for PY3. The analysis of Medicare savings based on *all LEJRs (IP+OP)* suggests that savings decreased substantially in the third year, primarily because mandatory CJR hospitals were less likely than other hospitals to provide TKA in the hospital outpatient setting. Savings are likely to continue to decline with more time under the OP TKA payment policy.

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

- **The relative decrease in average episode payments was driven by reductions in the use of institutional PAC.** Mandatory CJR hospitals discharged relatively fewer patients to an IRF (a 26.3% decrease from the CJR baseline proportion, $p<0.01$), resulting in a relative decrease in IRF payments of $539 per episode ($p<0.01$). CJR participant hospitals also discharged relatively fewer patients to a SNF (an 8.8% decrease from the CJR baseline proportion, $p<0.05$) and CJR patients with a SNF stay spent 2.5 fewer days in a SNF relative to control patients ($p<0.01$). These changes in SNF utilization resulted in a relative decrease in SNF payments of $935 per episode ($p<0.01$). Additionally, more CJR patients were first discharged to an HHA (a 21.9% increase from the CJR baseline proportion, $p<0.05$). The estimates presented here are based on only *inpatient LEJR episodes* and thus overestimate the impact; the estimates based on *all LEJRs (IP+OP)* are

---

11 Non-standardized paid amounts differ from the change in standardized allowed amounts used for the average episode payment impact analyses. In general, the change in paid amounts will be smaller because it does not include the change in beneficiary cost-sharing.

12 The $90.5 million is based on $109.4 million in reconciliation payments made to mandatory CJR hospitals less $18.9 million in repayments received from mandatory CJR hospitals.

in Appendix B (Exhibit B-3). While magnitudes differ, both sets of estimates indicate that the reduction in average episode payments was due to reductions in institutional PAC.

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

■ **Quality of care measures improved or were maintained under the CJR model.** During the first three performance years, the unplanned readmission rate decreased more for CJR episodes than for control group episodes, representing a 3.1% decrease from the CJR baseline (p<0.10). For elective LEJR episodes, there was a reduction in the complication rate, representing a 7.4% decrease from the baseline (p<0.01). There were no statistically significant changes in emergency department use or mortality. Further, there was no statistically significant change in the manipulation under anesthesia (MUA) rate for TKA episodes, a procedure that may be needed following TKA if the joint fails to achieve adequate range of motion. The estimates presented here are based on *inpatient LEJR episodes* and thus overestimate the impact; the estimates based on *all LEJRs (IP+OP)* are in Appendix B (Exhibit B-3). The estimates from both samples are similar in magnitude, but for estimates based on *all LEJRs (IP+OP)*, the reduction in the complication rate was not statistically significant.

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

■ **CJR and control patients who first received IRF or HHA care after their LEJRs showed similar improvement during the PAC stay, but results were mixed for those first receiving SNF care.** Beneficiaries first discharged to an IRF or HHA after their hospitalization showed similar short-term changes in functional status, measured as improvement in functional status and pain from the first to the last patient assessment completed during the PAC stay. For beneficiaries first discharged to a SNF, CJR patients had mixed changes in short-term functional status relative to control group patients.

■ **By the end of the episode, CJR and control survey respondents reported similar improvement in functional status and pain from before their episode.** On average, long-term functional recovery was similar between CJR and control patients who responded to our patient survey. When we analyzed results separately, based on where patients were first discharged after their hospital stay, our results suggest that functional recovery may have differed between CJR and control respondents depending on their first discharge setting. There is some evidence of worse functional improvement from before the episode to after the end of the episode among CJR respondents discharged from the hospital to an IRF, and to a lesser extent, CJR respondents first discharged to an HHA. However, our claims-based results indicate that the CJR model caused shifts of patients from more intense to less intense PAC settings, so these results by PAC setting may be biased downward (appearing to indicate worse CJR outcomes) if there were changes in patient complexity in each setting that we are unable to control for in our analyses.

■ **CJR and control respondents reported similar satisfaction with their overall recovery and care management.** However, CJR survey respondents, particularly those

first discharged to an IRF or SNF, reported less satisfactory care transition experiences and required more caregiver help at home than control respondents.

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

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

■ **The CJR model had no statistically significant impact on the total volume of elective LEJR discharges (inpatient plus outpatient).** The estimated impact of the CJR model on market-level LEJR discharge rates was a decrease of 0.06 discharges per 1,000 fee-for-service (FFS) beneficiaries, an estimate that is small and not statistically significant (p=0.31).

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

■ **Mandatory CJR hospitals had a decrease in patient complexity relative to control group hospitals.** A relative reduction in patient complexity could make it easier for mandatory CJR hospitals to receive reconciliation payments without improving the efficiency of how they provide care during the episode or improving the quality of care. Further, although we employed a rigorous risk adjustment methodology in our impact analyses, there may still be unmeasurable changes in patient complexity. If that is the case, the estimated reduction in average episode payments would be too high because it may, in part, be due to the relative reduction in the complexity of CJR patients.

■ **Greater relative complexity for CJR patients was suggested by characteristics that cannot be changed by hospital efforts to maximize patient outcomes.** Although mandatory CJR hospitals may try to reduce episode payments by supporting patients in changing behaviors prior to their LEJR that would improve their outcomes, the relative differences in patient characteristics between the CJR and control group patients does not indicate that this occurred. For example, for elective LEJRs without major complications or comorbidities, there were relative reductions in the proportion of CJR patients who were Black or eligible for Medicaid, among other changes. Changes in the proportion of Black or dually eligible patients raise the additional concern that existing disparities in receiving hip or knee replacements may be exacerbated under the CJR model.[13] We will continue studying this potential impact of the model.

■ **For the more complex episode types – elective LEJR with major complications or comorbidities and LEJRs due to fracture – the CJR patient population was not relatively healthier in the intervention period.**

---

[13]   For example, see Skinner J, Zhou W, Weinstein J. The influence of income and race on total knee arthroplasty in the United States. *J Bone Joint Surg Am*. 2006;88(10):2159-2166. doi:10.2106/JBJS.E.00271 and Chen J, Rizzo JA, Parasuraman S, Gunnarsson C. Racial disparities in receiving total hip/knee replacement surgery: the effect of hospital admission sources. *J Health Care Poor Underserved*. 2013;24(1):135-151. doi:10.1353/hpu.2013.0026

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

■ **We cannot draw a conclusion about the impact of the CJR model on payments for services provided during the 30 days following the episode.** Our estimate indicates that during the first three performance years the CJR model had no statistically significant impact on payments for services provided during the 30 days following the episode, which we monitor for any indications that services were postponed until after the end of the episode. However, during the baseline, there were differences in payment trends for services in the 30 days following the episode between CJR and control group hospitals, so we cannot tell if the lack of statistically significant impact was due to no relative change in payments after the episode or whether it was due to different underlying historical trends in payments.

3. *Financial risk or opportunity*

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

■ **In PY3, a lower proportion of CJR participant hospitals earned reconciliation payments than in previous years, which is consistent with model design changes that base a larger share of the target price on the regional average episode payments.** In PY3, 52% of hospitals earned reconciliation payments, compared to 58% in PY1 and 69% in PY2. The target price was more heavily based on historical hospital-specific episode payments in PY1 and PY2 to allow hospitals time to implement changes. In PY3, the quality-adjusted target price for 94% of mandatory CJR hospitals declined. As a result, they needed to make larger reductions in episode payments to earn a reconciliation payment.

■ **The majority of hospitals received reconciliation payments under this mandatory model, but hospitals with certain characteristics earned higher payments.** Mandatory CJR hospitals with a range of characteristics received reconciliation payments, but average reconciliation payments per episode were higher for mandatory CJR hospitals that had higher composite quality scores, had higher LEJR volume, were not-for-profit, and served less complex patient populations. The relationship between higher quality performance and higher reconciliation payments is consistent with model intent, to reward hospitals that focus on quality. The relationship between patient complexity and reconciliation payments may indicate that the simple risk stratification methodology used by CMS to set target prices based on Medicare Severity-Diagnosis Related Group (MS-DRG) and fracture status did not fully account for variations in patient complexity that affect episode payments within the four episode categories.

4. *Resources and market conditions*

a. *How did hospital resources and market conditions influence hospitals and other providers?*

■ **Interviewees indicated that financial arrangements between hospitals and surgeons, as well as the availability of surgeons and quality PAC providers in a market,**

affected the amount of control CJR participant hospitals had over care redesign across the LEJR episode. Hospitals that employed surgeons reported more success implementing new care redesign initiatives. Interviewees indicated that hospitals had less control over care redesign, and thus episode payments and quality, when surgeons performed LEJR at multiple hospitals. Interviewees also described challenges associated with the low supply of quality PAC providers, noting it posed a barrier to collaboration and limited the hospital's ability to influence PAC provider behavior.

5. *CJR participant hospital responses*

a. *How did CJR participant hospitals respond to Medicare coverage of TKA in the hospital outpatient setting?*

■ **Hospital representative interviewees reported confusion about how to interpret the TKA policy change and described a variety of strategies for responding to it.** For instance, some mandatory CJR hospitals reported establishing a default inpatient or outpatient status for all TKA patients. Others modified or created new algorithms, intake forms, or other tools to determine the appropriate setting for the LEJR patient, as well as enhanced internal documentation and auditing processes. Interviewees reported that while patient care was the same regardless of inpatient or outpatient status, the need for education for outpatients increased. Interviewees also voiced concerns about the negative impact of the rule change on measures of patient satisfaction, hospital payments, and reconciliation payments.

b. *How did CJR participant hospitals change care coordination?*

■ **Hospital representative interviewees described coordinating care for LEJR patients prior to the surgery, during the hospitalization for the LEJR, and after discharge from the hospital.** Although care coordination activities were not consistently attributed to the CJR model, interviewees noted that key coordination activities included patient education, discharge planning, risk stratification, data sharing, collaboration with PAC providers, and patient tracking and follow-up. Most interviewees discussed the importance of taking an interdisciplinary approach to care coordination, indicating different members of a care coordination team collaborated throughout the care pathway. Coordination activities generally did not vary for CJR and other LEJR patients, rather, changes were implemented for all LEJR patients, and CJR participant hospitals tailored care coordination approaches based on patients' needs. Interviewees indicated that care coordination typically required significant resources, and as a result some hospitals hired additional staff while others distributed responsibilities across existing roles.

c. *How did the CJR model affect patient and caregiver experience?*

■ **Hospital and PAC provider interviewees noted patients and caregivers could face challenges with direct discharge home or shorter SNF stays and described strategies to mitigate these challenges.** Challenges included patient push-back about discharge destination, increased reliance on caregivers, and the possibility that patients incurred

additional out-of-pocket expenses due to shifts in discharge destination. To address these challenges, interviewees discussed setting patient and caregiver expectations during pre-surgical education and supporting return to function through earlier ambulation and improved pain management protocols.

C. Discussion

This third annual evaluation report on the CJR model provides additional evidence that episode-based payments and quality measurement for LEJR can lower payments while improving or maintaining quality of care. Hospitals responded quickly to the incentives of the CJR model to achieve payment reductions, primarily by reducing institutional PAC use. Even with the reductions in average payments, quality of care was preserved and there are indications that the CJR model may be reducing unplanned readmissions. Further, after accounting for net reconciliation payments, the payment reductions made by mandatory CJR hospitals resulted in Medicare savings during the first three performance years. Medicare savings declined substantially in the third year because mandatory CJR hospitals shifted a smaller portion of patients to the outpatient setting after Medicare's change in TKA coverage. Medicare savings for all hospitals that ever participated in the CJR model would likely differ from the savings from mandatory CJR hospitals presented in this report. CJR participant hospitals described strategies to shift post-acute care from more- to less-intensive settings to meet model objectives. However, the ability to engage providers along the care continuum in efforts to reduce payments varied across participant hospitals and across markets. This evaluation indicates that a range of hospitals, with a variety of resources and circumstances, can and do respond to the incentives under a mandatory model to reduce episode payments.

In future reports, we will deepen our understanding of the impact of the CJR model. We will be able to examine differences in market level effects of the model by comparing MSAs that remain fully mandatory with those that have a mixture of continuing and exiting hospitals. We will continue to assess unintended consequences on volume and the adoption of practice changes, and further explore the variation in patient complexity and its impact on reconciliation payments and Medicare savings. As the payment landscape continues to evolve, we will further refine our impact estimates on payments, Medicare program savings, and quality by considering other policies and models that affect service use and payments.

I.   Introduction

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

The third annual CJR model evaluation report presents findings from the first three performance years, which include episodes initiated on or after April 1, 2016 that ended by December 31, 2018. During the first two years, the model was mandatory for nearly all acute care hospitals in 67 geographic areas, defined by MSAs.[16] At the start of PY3, the number of mandatory MSAs was scaled back to the 34 MSAs with the highest historical payments; hospitals in the other 33 MSAs were given the opportunity to continue to participate voluntarily. This report focuses on the hospitals in the 34 MSAs that were required to participate in the CJR model throughout the intervention period (mandatory CJR hospitals).

A.   The CJR Model

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

---

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

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

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

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

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

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

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

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

Hospitals with LEJR episode payments *below* their quality-adjusted target price and an "acceptable" or higher composite quality score receive a reconciliation payment. The reconciliation payment equals the difference between the quality-adjusted target price and actual episode payments, up to a stop-gain limit. Starting in PY2, hospitals with episode payments *above* their quality-adjusted target price repay Medicare the difference, subject to a stop-loss limit. In PY1, this repayment responsibility was forgiven to allow hospitals time to gain experience under the CJR model before implementation of two-sided risk. Both opportunity and risk increase over time as stop-gain and stop-loss limits increase. The stop-gain limit increased from 5% in PY1 to 10% in PY3 and the stop-loss limit increased from 0% to 10% over the same period.

Mandatory, randomized design of the original CJR model. The original mandatory, randomized design of the CJR model resulted in a diverse group of CJR participant hospitals,

---

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

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

Changes to the CJR model in 2018. Effective January 2018, CMS scaled back the CJR model to the 34 MSAs with the highest historical episode payments (mandatory MSAs). Hospitals in these mandatory MSAs that were not designated as low-volume or rural were required to continue their participation in the CJR model. The number of hospitals required to participate was reduced from 800 at the end of PY2 to 378. In January 2018, CMS allowed hospitals in the 33 lower payment MSAs (voluntary MSAs) and all hospitals that were designated as low-volume or rural hospitals a one-time opportunity to opt-in to the CJR model for the PY3-5 (Exhibit 1).

---

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

Exhibit 1:   Location of mandatory and voluntary CJR MSAs in PY3

Always mandatory CJR MSAs (n=34)   Voluntary CJR MSAs (n=33)
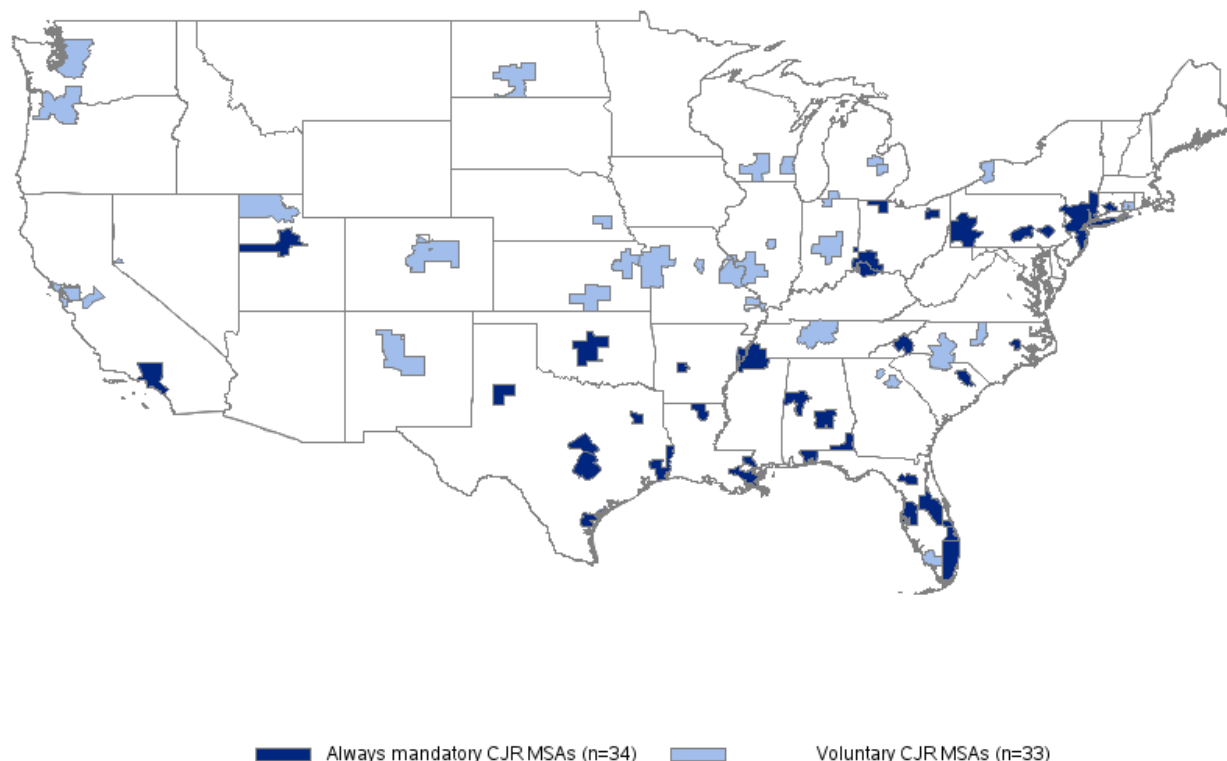
Source: Final Rule. Medicare Program; Cancellation of Advancing Care Coordination Through Episode Payment and Cardiac Rehabilitation Incentive Payment Models; Changes to Comprehensive Care for Joint Replacement Payment Model: Extreme and Uncontrollable Circumstances Policy for the Comprehensive Care for Joint Replacement Payment Model. December 1, 2017. Available at: https://www.federalregister.gov/documents/2017/12/01/2017-25979/medicare-program-cancellation-of-advancing-care-coordination-through-episode-payment-and-cardiac

Note: MSA = metropolitan statistical area, PY = performance year.

The randomized design of the CJR model was not completely preserved when the model was scaled back in PY3 because the 34 mandatory MSAs were identified using a median split based on historical payments rather than the original sampling strata that categorized MSAs as high payment. Most, but not all, of the MSAs in the top two payment quartiles were categorized in the mandatory MSA group. The characteristics of mandatory CJR hospitals reflect their location in higher payment markets; however, the mandatory CJR hospitals remain a diverse group with variation in terms of LEJR volume, patient complexity, and institutional characteristics (Section II.A.1). Hospitals in high-payment MSAs that were eligible but not selected for the CJR model serve as the control group for mandatory CJR hospitals.

Other policy changes in 2018. CMS implemented broad changes to Medicare coverage that affect the evaluation of the CJR model. The CJR model implemented episode-based payments for inpatient LEJR, which are primarily hip replacements (THA) and knee replacements (TKA). Effective January 2018, CMS removed TKA from Medicare's inpatient only list to allow Medicare coverage of TKAs performed in the hospital outpatient department. When TKA is performed on an inpatient, the hospital's payment is made under the IPPS and the surgery triggers an episode if the

hospital is participating in the CJR model. When TKA is performed on an outpatient, the hospital's payment is made under the outpatient payment system and the surgery does not trigger an episode. Although this policy change was independent of the CJR model, CJR participant hospitals may have responded to it differently than non-CJR hospitals. Section II.A.2 describes how CJR participant hospitals responded to the policy and how the policy affects our analysis of the impact of the CJR model on average episode payments and Medicare savings.

Additional detail about the CJR model is available in the Second Annual Evaluation Report, available at https://innovation.cms.gov/Files/reports/cjr-secondannrpt.pdf, and the CMMI webpage, https://innovation.cms.gov/initiatives/cjr.

B. Evaluation Conceptual Framework

The conceptual framework for the evaluation of the CJR model (Exhibit 2) reflects the fundamental features of the model and is informed by health services research literature, including evaluations of other episode-based payment approaches.[20] The evaluation framework focuses on the hospital where the LEJR episode begins because the hospital has the incentives to control payments and improve quality across the entire episode. The hospital's resources and market conditions will affect its decisions about whether and how to respond to the model.

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

---

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

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

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

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

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

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

- Choice of response
- Care redesign

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

- Reconciliation payments
- Perceptions of financial pressure

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

- Relationships and resources

Note: LEJR = lower extremity joint replacement.

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

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

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

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

Choice of response. Hospital leaders must consider multiple organizational factors, in addition to the potential for financial risk or opportunity, and internal and external resources, in making the business case for whether and how to respond to the CJR model. Orthopedic surgery is one of multiple service lines that compete for staff and other resources. The CJR model is one initiative that may or may not align with initiatives from other payers, state-specific policies, local labor markets, and other factors. Site visits and structured interviews provide data about how hospitals made decisions about their response to the model. (Section II.E.1 examines how participant hospitals changed care coordination.)

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. Because of the changes in the calculation of the quality-adjusted target price, the amount of financial pressure on hospitals will shift over time. (Section II.C.1 explores the market, hospital, and patient characteristics associated with the amount of reconciliation payments.)

Resources and market conditions. A hospital's internal resources and market conditions will provide opportunities or constraints on its responses to the model. Hospitals with more capital and operational resources, such as dedicated care coordination staff or robust health information technology infrastructure, may be better situated to redesign care for LEJR episodes. Other hospital resources – such as leadership support, experience with episode-based payment or similar payment models, ownership of PAC providers, or employment of surgeons – may also affect their choices as well as their success in reducing payments below their quality-adjusted target price. Market conditions, such as the supply and characteristics of other providers involved in the episode, will affect how and whether hospitals garner support for delivering care more efficiently during the episode. 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.D.1 examines how hospital resources and market conditions influenced hospitals and other providers; detailed case studies can be found in the Provider Experiences Report.)

II. Results

A. Study Population and Policy Context

1. *How did the characteristics of CJR participant hospitals compare with those of other hospitals?*

This evaluation report focuses on the 378 hospitals that were required to continue participating in the CJR model in PY3. Throughout the report, we refer to these as "mandatory CJR hospitals." This section describes the mandatory CJR hospitals and their markets. We compared their characteristics to those of all other IPPS hospitals and markets in the U.S. We also compared characteristics of the CJR participant hospitals in voluntary MSAs that opted-in to those that did not opt-in (non-opt-in) and identified factors associated with the decision to continue participation.

a. Key findings

- ■ Mandatory CJR hospitals were a diverse group of hospitals that vary in terms of size, LEJR experience, market conditions, episode payments, and patient populations. Mandatory CJR hospitals were similar to all other IPPS hospitals in the U.S. on many characteristics.

- ■ Compared to all other IPPS hospitals, mandatory CJR hospitals had a similar annual LEJR volume, but had historically higher average episode payments and discharged a greater proportion of patients to institutional PAC settings. These differences are generally explained by CMS' sampling strategy, which targeted higher payment, urban MSAs for participation in the CJR model.

- ■ In voluntary MSAs, opt-in hospitals were more likely to receive reconciliation payments in the first two years of the model than non-opt-in hospitals.

b. Methods

We calculated summary statistics (frequencies, means, and ranges) of key variables to describe mandatory CJR hospitals. Comparisons were made across groups (mandatory CJR vs. all other IPPS hospitals, opt-in vs. non-opt-in hospitals in voluntary MSAs) using chi-square tests for categorical variables and t-tests for continuous variables.[21]

c. Results

Mandatory CJR hospitals

Mandatory CJR hospitals were distributed across the country, although they were more concentrated in the Northeast and South. This is because the CJR model focused on MSAs with

---

[21] Appendix D – Definition of MSA, Hospital, and Patient Characteristics

higher historical payments and the Northeast and South are typically higher cost regions (Exhibit 3).

Exhibit 3:   Mandatory CJR hospitals distributed across the U.S. with greater representation in the Northeast and South
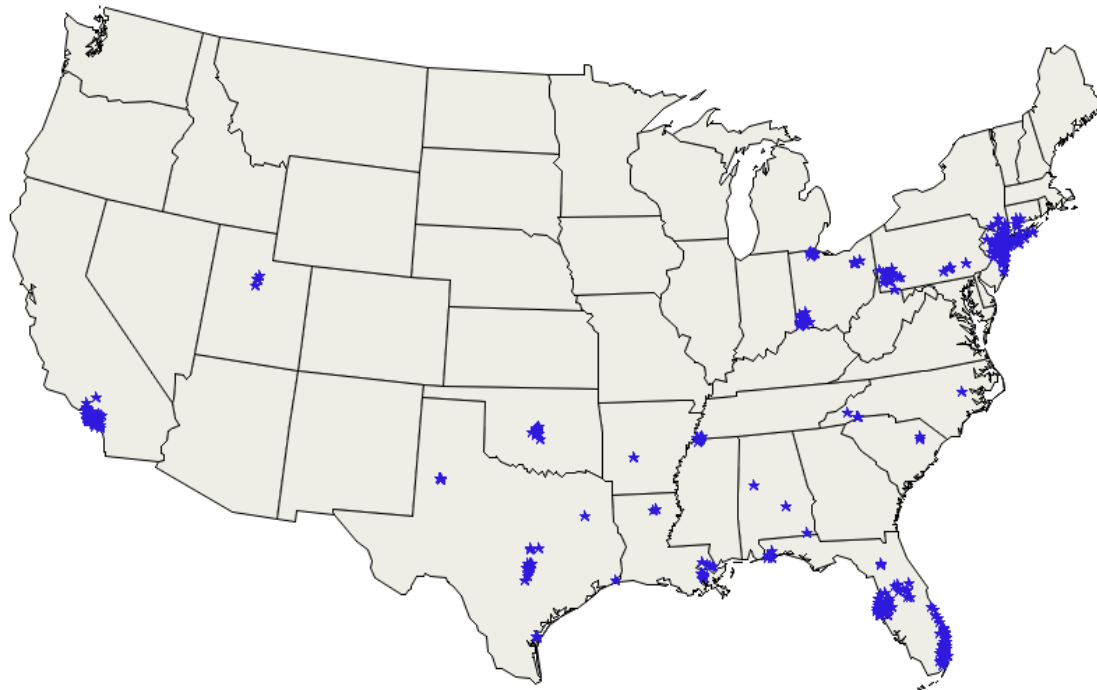
Source: CJR evaluation team analysis of the CJR participant list (October 2019) and Medicare claims and enrollment data for episodes initiated in 2012 through 2014 that ended between April 2012 and March 2015 (baseline) and episodes initiated on or after April 2016 and ended by December 2018 (intervention).

Notes: Mandatory CJR hospitals include hospitals with any Medicare LEJR patients during the baseline and intervention periods. LEJR = lower extremity joint replacement.

Because of the mandatory, randomized design, a spectrum of hospitals with varying levels of infrastructure, LEJR experience, market resources, care redesign experience, episode payments, and patient populations participated. Sixty percent of mandatory CJR hospitals were not for profit, 28% were for profit, and 13% were government owned. Forty-three percent of mandatory CJR hospitals had any affiliation with a medical school. Mandatory CJR hospitals varied in size, with an average bed count of 277, ranging from 8 to 1,928. Hospitals also ranged in terms of annual baseline LEJR volume (mean: 157; range: 8 - 1,457), LEJR as a percent of total baseline discharges (mean: 6.1%; range: 0.3% - 86.9%), and baseline LEJR market share (mean: 7.4%; range: 0.04% - 100.0%). Ten percent of mandatory CJR hospitals previously participated in the Bundled Payments for Care Improvement (BPCI) initiative for any clinical episode. Mandatory CJR hospitals had a wide range in average baseline episode payments (mean: $32,516; range:

$19,268 - $66,074) and treated diverse patient populations; the average baseline hierarchical condition category (HCC) score of their LEJR patients was 1.64, ranging from 0.79 to 3.22.[22]

Mandatory CJR hospitals compared with all other U.S. hospitals

We compared mandatory CJR hospitals to all other IPPS hospitals in the U.S. that perform LEJRs. The purpose of this analysis was to understand the characteristics of both groups of hospitals; this analysis did not assess the generalizability of our results to the larger pool of acute care hospitals in the U.S. Future analyses under this evaluation will assess generalizability.

Mandatory CJR hospitals were similar to all other IPPS hospitals in the U.S. when comparing ranges on a variety of characteristics. This suggests that a wide spectrum of IPPS hospitals were included in the mandatory group. For example, average baseline episode payments ranged from $19,268 to $66,074 for mandatory CJR hospitals, and from $12,447 to $78,924 for all other IPPS hospitals (Exhibit 4). The range for all other IPPS hospitals includes the range for mandatory CJR hospitals.

---

[22] The HCC score predicts Medicare spending in the coming year for a given Medicare beneficiary compared to the average beneficiary. Beneficiaries with scores above 1.0 are more complex and predicted to have higher costs in the coming year than the average beneficiary and beneficiaries with scores below 1.0 are healthier and predicted to have lower costs than the average.

Exhibit 4:   Across a number of characteristics, ranges in values for mandatory CJR hospitals overlap with the ranges for all other IPPS hospitals

Hospital characteristics
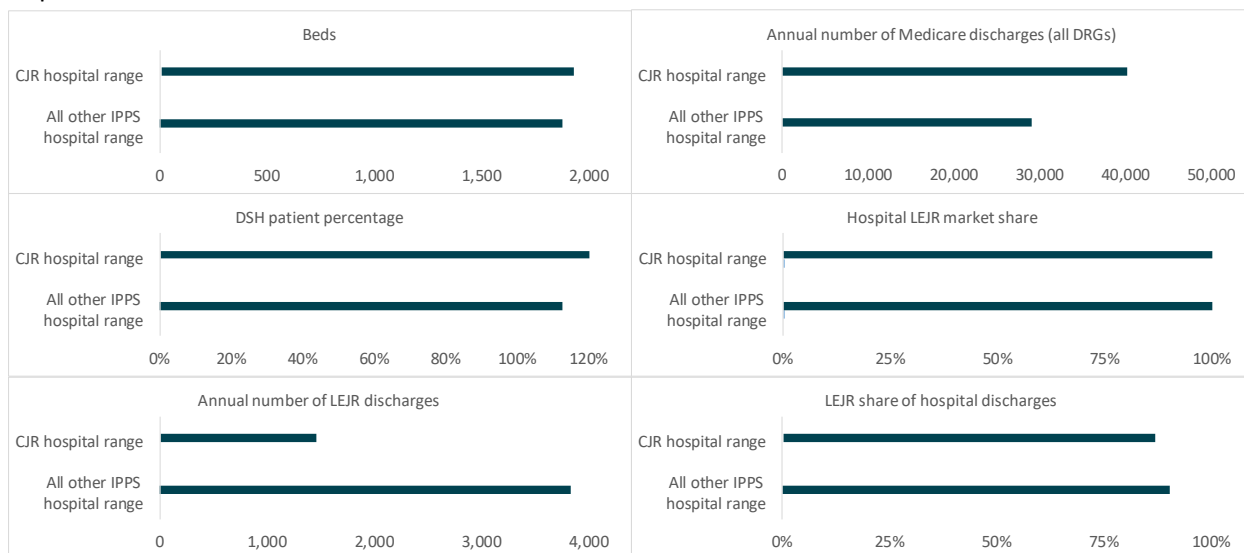
LEJR episode characteristics
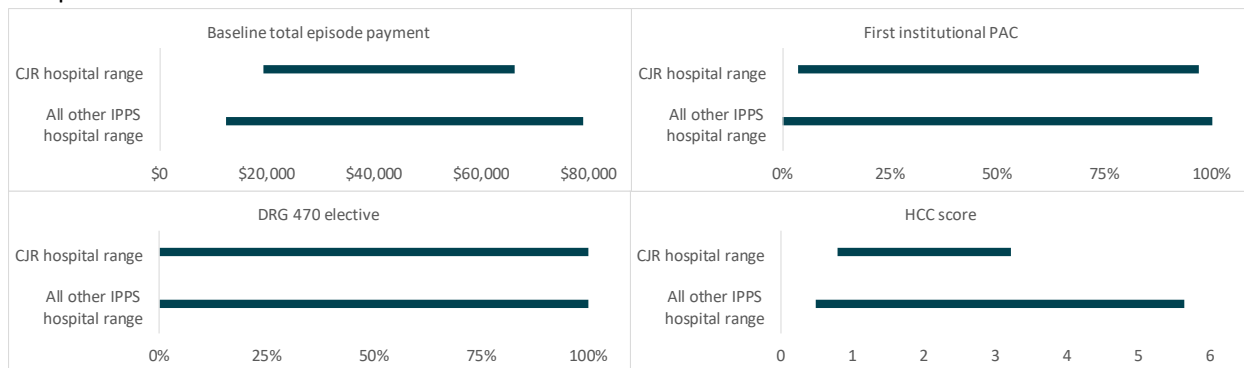
Source: CJR evaluation team analysis of December 2016 POS, FY 2016 CMS Annual IPPS, and Medicare claims and enrollment data for episodes initiated in 2012 through 2014 that ended between April 2012 and March 2015.

Notes: Mandatory CJR hospitals are defined as all hospitals located in CJR-participating MSAs and required to participate in the CJR model in PY3. All other IPPS hospitals were located across the U.S., had at least one LEJR during the baseline period, and were not participating in the CJR model in PY3.

Baseline total episode payments are standardized allowed payments and are not risk-adjusted.

DRG = diagnosis related group, DSH = disproportionate share hospital, FY = fiscal year, HCC = hierarchical condition category, IPPS = inpatient prospective payment system, LEJR = lower extremity joint replacement, MSA = metropolitan statistical area, PAC = post-acute care, POS = provider of services, PY = performance year.

While the ranges of values for mandatory CJR hospitals were generally similar to all other IPPS hospitals in the U.S., it is important to note that there are some differences between the two groups when comparing averages or proportions. These differences are largely driven by the design of the model, namely the implementation of the model in populous urban areas with historically high episode payments. CJR MSAs were randomly sampled from eligible MSAs that had an urban core population of at least 50,000 and at least 400 LEJR cases annually. Further, CJR MSAs are concentrated in areas that historically had the highest average LEJR payments.

Mandatory CJR hospitals were more likely to be for profit (27.8% vs. 20.2%, p<0.01) and located in the Northeast (31.5% vs. 14.2%) and South (44.7% vs. 38.5%, p<0.01) than all other IPPS hospitals (Exhibits 5a and 5b). All mandatory CJR hospitals were located in urban areas, compared with 94.9% of all other hospitals paid under Medicare's IPPS, and other differences between the CJR participant hospitals and all other IPPS hospitals reflected this urban focus. Mandatory CJR hospitals were larger hospitals with higher average bed counts (277 vs. 200, p<0.01) and higher average annual Medicare discharges (4,145 vs. 3,216, p<0.01) than other IPPS hospitals across the country. On average, they served a greater proportion of low income patients (disproportionate share hospital (DSH) patient percentage: 32.3% vs. 27.7%, p<0.01) and had lower average LEJR market share (7.4% vs. 31.9%, p<0.01). Mandatory CJR hospitals were also more likely to have an affiliation with a medical school than all other IPPS hospitals (42.9% vs. 34.5%, p<0.01).

Exhibit 5a & 5b: Mandatory CJR hospitals differed from all other IPPS hospitals on characteristics related to the CJR model design

Baseline characteristic		Mandatory CJR hospitals (N=378)	All other IPPS hospitals (N=2,536)	p-value
		%	%	
Ownership	Non-profit	59.5%	63.8%	p<0.01
	For-profit	27.8%	20.2%	
	Government	12.7%	15.9%	
Census region	Northeast	31.5%	14.2%	p<0.01
	South[a]	44.7%	38.5%	
	Midwest	6.6%	27.0%	
	West	17.2%	20.3%	
Urban	Located in an urban area	100.0%	94.9%	p<0.01
Teaching hospital	Yes	42.9%	34.5%	p<0.01
Prior BPCI experience	Ever participated in the BPCI initiative (LEJR or non-LEJR)	10.3%	12.5%	p=0.23

Baseline characteristic[b]	Mandatory CJR hospitals (N=378)	All other IPPS hospitals (N=2,536)	p-value
	Mean	Mean	
Annual number of LEJR discharges	157	158	p=0.96
Total payment for inpatient stay and 90 day post-discharge period[c]	$32,516	$28,005	p<0.01
HCC score	1.64	1.43	p<0.01
Percent MS-DRG 470 elective	71.2%	75.5%	p<0.01
First PAC in an institutional setting	60.5%	50.4%	p<0.01
Bed count	277	200	p<0.01
Annual number of Medicare discharges (all MS-DRGs)	4,145	3,216	p<0.01
DSH patient percentage	32.3%	27.7%	p<0.01
LEJR share of hospital discharges	6.1%	6.8%	p=0.29
Percent of MSA's LEJRs performed at hospital	7.4%	31.9%	p<0.01

Source: CJR evaluation team analysis of December 2016 POS, FY 2016 CMS Annual IPPS, Bundled Payments for Care Improvement initiative participant list, and Medicare claims and enrollment data for episodes initiated in 2012 through 2014 that ended between April 2012 and March 2015.

Notes: Mandatory CJR hospitals are defined as all hospitals located in CJR-participating MSAs and required to participate in the CJR model in PY3. All other IPPS hospitals were located across the U.S., had at least one LEJR during the baseline period, and were not participating in the CJR model in PY3.

Differences tested using chi-square tests for categorical variables and t-tests for continuous variable. Differences that are significant at the 99%, 95%, or 90% significance level are indicated by red, orange, or yellow shaded cells, respectively.

BPCI = Bundled Payment for Care Improvement, DSH = disproportionate share hospital, FY = fiscal year, HCC = hierarchical condition category, IPPS = inpatient prospective payment system, LEJR = lower extremity joint replacement, MSA = metropolitan statistical area, MS-DRG = Medicare Severity-Diagnosis Related Group, PAC = post-acute care, POS = provider of services, PY = performance year.

[a]   Hospitals located in Puerto Rico are included in the South.

[b]   The baseline outcomes are not risk adjusted.

[c]   Standardized allowed payments.

By design, mandatory CJR hospitals had higher average episode payments at baseline than all other IPPS hospitals ($32,516 vs. $28,005, p<0.01). Consistent with higher episode payments, mandatory CJR hospitals served more complex patient populations, as illustrated by the higher average HCC score (1.64 vs. 1.43, p<0.01) and lower percent of episodes that were MS-DRG 470 elective (71.2% vs. 75.5%, p<0.01). Also, patients treated by mandatory CJR hospitals were more likely to be discharged to an institutional PAC setting (60.5% vs. 50.4%, p<0.01).

Mandatory CJR MSAs compared with all other MSAs in the U.S.

We also compared characteristics of the 34 mandatory CJR MSAs with all other MSAs. The mandatory CJR MSAs had larger populations, on average, than all other MSAs (an average population of 1.8 million vs. approximately 590,000, p<0.10, Exhibit 6). CMS randomly sampled CJR MSAs, and by chance, the two largest MSAs were selected to participate in the CJR model (New York-Newark-Jersey City, NY-NJ-PA and Los Angeles-Long Beach-Anaheim, CA). Further, the CJR MSAs had a larger average population because the CJR model excluded smaller MSAs with less than 400 annual LEJRs (n=156).

Consistent with larger urban areas and the sampling criteria mentioned above, the mandatory CJR MSAs had more hospitals performing LEJRs in the market (p<0.10), and more LEJRs performed annually (p<0.10). On average, mandatory CJR MSAs and all other MSAs were both characterized as highly concentrated markets as measured by the Herfindahl-Hirschman Index (HHI). However, the mandatory CJR MSAs were more competitive with a significantly lower HHI (3,675 vs. 5,157, p<0.01).[23]

Exhibit 6: **Mandatory CJR MSAs had larger populations, more hospitals, and more LEJR episodes than all other MSAs in the U.S.**

Baseline MSA characteristic	Mandatory CJR MSAs (N=34)		All other MSAs (N=354)		p-value
	Mean (range)	Median (IQR)	Mean (range)	Median (IQR)	
Population size	1,840,635 (96,804 − 19,865,045)	449,934 (268,707 − 1,337,014)	589,705 (54,634 − 9,516,448)	213,790 (135,240 − 518,387)	p<0.10
Number of hospitals performing LEJRs	15 (1 − 127)	4 (3 − 18)	5 (1 − 82)	3 (2 − 5)	p<0.10
Annual number of LEJRs	2,278 (435 − 22,104)	1,078 (587 − 2,217)	994 (8 − 15,385)	508 (276 − 1,052)	p<0.10
HHI for LEJR at ACH	3,675 (422 − 10,000)	3,382 (1,606 − 5,063)	5,430 (269 − 10,000)	5,157 (3,141 − 7,428)	p<0.01

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

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

All other MSAs includes MSAs that previously participated in the CJR model (33 voluntary MSAs).

---

[23] The HHI is defined as the sum of the squares of the market shares of each hospital performing LEJR in the MSA. The HHI ranges from close to zero (indicating nearly perfect competition) to 10,000 (indicating a monopoly). The U.S. Department of Justice considers a market with an HHI of 2,500 or greater to be a highly concentrated market.

ACH = acute care hospitals, HHI = Herfindahl-Hirschman Index, IQR = interquartile range, LEJR = lower extremity joint replacement, MSA = metropolitan statistical area.

Opt-in hospitals in voluntary MSAs compared with non-opt-in hospitals in voluntary MSAs

In January 2018, CMS made participation voluntary for hospitals in the 33 original CJR MSAs with the lowest historical episode payments and provided hospitals with a one-time opportunity to continue participating in the CJR model. Seventy-six hospitals opted to continue participation in the CJR model and 233 hospitals did not opt-in. The voluntary MSAs were concentrated in the Midwest and West (Exhibit 7). A higher percent of opt-in hospitals in voluntary MSAs were located in the West, specifically the Pacific Northwest.

Exhibit 7: **Opt-in and non-opt-in hospitals in voluntary MSAs were distributed across the U.S. with opt-in hospitals more concentrated in the West**
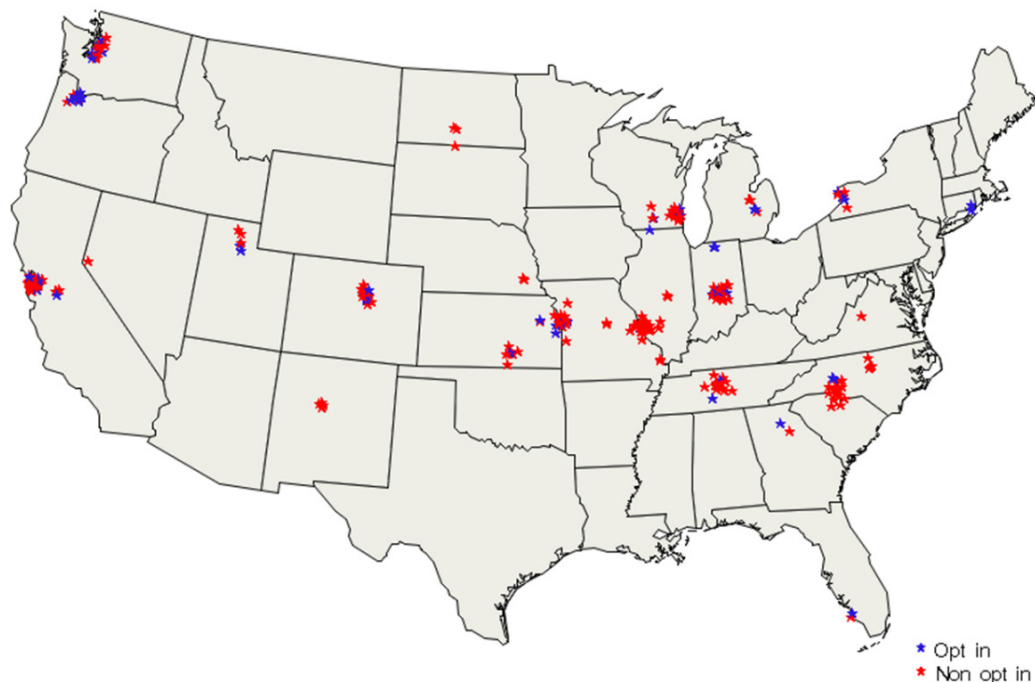
Source: CJR evaluation team analysis of the CJR participant list (October 2019) and CJR participants prior to February 2018 (available at https://innovation.cms.gov/Files/x/cjr-hospitallist-pre0218.xlsx).

Note: MSA = metropolitan statistical area.

Opt-in hospitals financially benefited from participation in the CJR model in the first two performance years and this likely influenced their decision to opt-in in PY3. Three-quarters of opt-in hospitals earned reconciliation payments in PY1 and 90.7% earned reconciliation payments in PY2. This was significantly higher than the proportion of non-opt-in hospitals that earned reconciliation payments (46.6% in PY1 and 66.0% in PY2; both $p<0.01$, Exhibit 8a & 8b).

Consistent with the increased likelihood of receiveing reconciliation payments, a higher percent of opt-in hospitals started the CJR model with historical episode payments below their PY1 quality-adjusted target price than non-opt-in hospitals (73.3% vs. 45.8%; $p<0.01$). Average baseline

episode payments were also significantly lower for opt-in hospitals than for non-opt-in hospitals ($23,938 vs. $26,470, p<0.01). Their lower use of institional PAC may partially explain these lower payments. On average, opt-in hospitals discharged fewer patients to institutional PAC settings at baseline (40.7% vs. 45.0%, p<0.01).

Exhibit 8a & 8b: Opt-in and non-opt-in hospitals in voluntary CJR MSAs differed in terms of LEJR-related characteristics and financial performance in PY1 and PY2

Baseline characteristic		Opt-in hospitals in voluntary CJR MSAs (N=76)	Non-opt-in hospitals in voluntary CJR MSAs (N=233)	p-value
		%	%	p-value
Ownership	Non-profit	72.4%	64.0%	p=0.36
	For-profit	14.5%	16.7%	
	Government	13.2%	19.3%	
Census region	Northeast	7.9%	1.7%	p<0.01
	South	11.8%	16.3%	
	Midwest	31.6%	48.1%	
	West	48.7%	33.9%	
Teaching hospital	Yes	36.8%	39.1%	p=0.73
Prior BPCI experience	Ever participated in the BPCI initiative (LEJR or non-LEJR)	11.8%	8.6%	p=0.40
Historical episode payments relative to PY1 quality-adjusted target price	Below PY1 target price	73.3%	45.8%	p<0.01
Earned reconciliation payments	PY1	75.0%	46.6%	p<0.01
	PY2	90.7%	66.0%	p<0.01

Baseline characteristic[a]	Opt-in hospitals in voluntary CJR MSAs (N=76)	Non-opt-in hospitals in voluntary CJR MSAs (N=233)	
	Mean	Mean	p-value
Annual number of LEJR discharges	264	163	p<0.01
Total payment for inpatient stay and 90 day post-discharge period[b]	$23,938	$26,470	p<0.01
HCC score	1.23	1.35	p<0.01
Percent MS-DRG 470 elective	85.2%	78.5%	p<0.01
First discharged to institutional PAC setting	40.7%	45.0%	p<0.10
Bed count	219	207	p=0.62
Annual number of Medicare discharges (all MS-DRGs)	3,520	3,103	p=0.28
DSH patient percentage	21.6%	26.9%	p<0.01
LEJR share of hospital discharges	13.2%	5.8%	p<0.01
PY1 quality composite score	13.1	10.4	p<0.01
PY2 quality composite score	12.8	10.3	p<0.01
Percent of MSA's LEJR performed at hospital	15.9%	8.0%	p<0.01

Source: CJR evaluation team analysis of December 2016 POS, FY 2016 CMS Annual IPPS, Bundled Payments for Care Improvement initiative participant list, and Medicare claims and enrollment data for episodes initiated in 2012 through 2014 that ended between April 2012 and March 2015.

Notes: Opt-in hospitals in voluntary MSAs chose to continue participation in the CJR model in PY3. Non-opt-in hospitals in voluntary MSAs did not opt to continue participation.

Differences tested using chi-square tests for categorical variables and t-tests for continuous variable. Significant at the 99%, 95%, or 90% significance level are indicated by red, orange, or yellow shaded cells, respectively.

BPCI = Bundled Payments for Care Improvement, DSH = disproportionate share hospital, FY = fiscal year, HCC = hierarchical condition category, IPPS = Inpatient Prospective Payment System, LEJR = lower extremity joint replacement, MSA = metropolitan statistical area, MS-DRG = Medicare Severity-Diagnosis Related Group, PAC = post-acute care, POS = provider of services, PY = performance year.

[a] The baseline outcomes are not risk adjusted.

[b] Standardized allowed payments.

Further, opt-in hospitals had significantly higher average quality scores in the first two years of the model (13.1 and 12.8) than non-opt-in hospitals (10.4 and 10.3; p<0.01). Under the CJR model, the target prices are quality-adjusted, meaning hospitals with higher quality scores can earn higher reconciliation payments or have less repayment responsibility to CMS.

The decision to opt-in to the CJR model in PY3 was also correlated with hospital LEJR volume and importance of LEJR to the hospital service line. Opt-in hospitals had higher LEJR volume, performing an average of 264 LEJRs annually, compared to 163 LEJRs annually for non-opt-in hospitals (p<0.01). They also had higher LEJR share of total discharges (13.2%) and higher LEJR market share (15.9%) than the non-opt-in hospitals (5.8% and 8.0%, respectively; both p<0.01).

Opt-in hospitals served less complex patient populations, on average. They had lower average HCC scores for their LEJR patients (1.23 vs. 1.35, p<0.01), performed a higher percent of LEJRs

for the less complicated MS-DRG 470 elective episode type (85.2% vs. 78.5%, p<0.01), and had lower average DSH patient percentages (21.6% vs. 26.9%, p<0.01) than non-opt-in hospitals.

Opt-in and non-opt-in hospitals in voluntary CJR MSAs were similar in terms of ownership, teaching status, past participation in the BPCI initiative (any clinical episode), bed count, and annual total number of Medicare discharges (all MS-DRGs).

d. Conclusion

Mandatory CJR hospitals were a diverse group of hospitals varying in terms of size, LEJR experience, market conditions, episode payments, and patient populations. Mandatory CJR hospitals were similar to all other IPPS hospitals in the U.S. on many characteristics. The ranges in characteristics for the two groups overlapped, indicating that a wide spectrum of IPPS hospitals are represented by mandatory CJR hospitals. However, there were some significant differences between mandatory CJR hospitals and all other IPPS hospitals. Mandatory CJR hospitals were more likely to be larger teaching hospitals located in the Northeast and South with higher average episode payments and more complex patient populations. These differences were likely driven by the design of the model, which was rolled out in metropolitan areas with large populations and a minimum of 400 LEJRs per year. In addition, only high-payment MSAs were required to continue participation in the CJR model in PY3.

These results provide information on factors associated with the decision of hospitals in voluntary MSAs to continue participation in the CJR model. The decision to opt-in was correlated with a variety of LEJR-related factors and financial performance in the first two performance years. Opt-in hospitals had higher LEJR volume and served less complex patient populations. The majority received reconciliation payments in the first two performance years of the model and nearly three-quarters started the CJR model with historical episode payments below the PY1 quality-adjusted target price. The non-opt-in hospitals likely chose not to participate in the CJR model because it did not financially benefit them.

2. *Did CJR participant hospitals respond differently than other hospitals to Medicare coverage of TKA in the hospital outpatient setting and how would this affect the evaluation of the model?*

CMS removed TKA from Medicare's inpatient only list, effective January 2018, to allow Medicare coverage of TKAs performed in the hospital outpatient department. As a result, Medicare will pay for TKAs performed as a hospital inpatient or outpatient procedure. When a TKA is performed in the inpatient setting, the hospital's payment is made under the IPPS and, if the hospital participates in the CJR model, the surgery triggers a CJR episode. When a TKA is performed in the hospital outpatient setting, the hospital's payment is made under the outpatient payment system and the surgery does not trigger a CJR episode.

In 2018, about 19% of TKAs were performed in the outpatient setting for mandatory CJR hospitals compared to about 29% for control group hospitals (p<0.01). In this section, we explore the factors associated with the choice of inpatient or outpatient setting. Further, we evaluate whether the CJR

model contributed to the smaller share of TKAs performed in the outpatient setting for CJR participant hospitals and how this could affect the evaluation of the CJR model.

---

a. Key Findings

- Mandatory CJR hospitals performed a smaller share of TKAs in the outpatient setting than control group hospitals (19% versus 29%, $p<0.01$).

- Evidence indicates that the difference in the share of TKAs performed in the outpatient setting between the CJR and control groups is due to the CJR model.

- Because of the difference in the shares of TKAs performed in the outpatient setting, impact estimates from an analysis including only inpatient LEJR episodes are biased and overestimate the impact of the CJR model on payment and service use reductions.

- Impact estimates based on all LEJRs, including outpatient TKAs, underestimate the impact of the CJR model because OP TKAs are not subject to CJR treatment, but are likely to be closer to the actual impact of the CJR model.

---

b. Methods

We evaluated the adoption of outpatient TKA after the change in the TKA coverage policy by calculating the share of TKAs performed in the outpatient setting for mandatory CJR hospitals and control group hospitals. We hypothesized that differences in OP TKA shares may be explained by patient complexity during the performance year and anchor hospitalization length of stay (LOS) prior to the change in TKA coverage policy. To examine these hypotheses, we calculated the average HCC score for TKA episodes in 2018 and average LOS in 2017 at the hospital-level and then estimated the average OP TKA share by categories of these measures. Differences in average OP TKA shares between CJR and control group hospitals were evaluated using t-tests. To isolate the effect of CJR participation from other factors that might cause differences in outpatient rates, we estimated risk-adjusted shares of TKA performed in the outpatient setting by adjusting for average hospital length of stay prior to the TKA policy change, as well as other hospital, regional, and patient characteristics.

More detail on these methods is available in Appendix E (Section IV).

c. Results

Factors that may affect whether a TKA is performed in the inpatient or outpatient setting

The decision of hospitals and physicians regarding whether to perform TKAs in the inpatient or outpatient setting is likely based on many factors:

- According to CMS guidance, the decision is to be based on the physician's clinical expectation of the patient's anticipated LOS in the acute care hospital setting and should consider the individual's unique clinical circumstances.[24]

- According to our interviews with hospital representatives, key factors that influence the decision include patient comorbidities, age, post-acute care needs, and social supports.[25] The expectation of physicians, patients, and caregivers about the need for a SNF stay following the TKA procedure would affect site of the surgery because patients who receive their TKA in the outpatient setting would not be eligible for Medicare coverage of a SNF stay.

- Clinical review panel members indicated that a hospital's average LOS for inpatient TKAs prior to the policy change could also affect the choice of setting. Members hypothesized that hospitals with historically short lengths of stay would have higher shares of outpatient TKAs than hospitals with long lengths of stay prior to the TKA policy change.

- The hospital's capacity to perform surgeries on an outpatient basis may affect the decision.

- OP TKA shares of CJR and control group hospitals may reflect practice patterns in their geographic areas with respect to tendencies to perform TKAs in the outpatient setting.

Financial incentives could also influence the decision:

- Medicare pays hospitals more for inpatient TKAs than for outpatient TKAs so, depending on the hospital costs for providing the procedure in the different settings, the hospital margin may differ by setting.[26]

- CJR participant hospitals have an additional financial incentive to perform TKAs in the inpatient setting:[27]

  - Patients who could appropriately receive TKAs as outpatients would be expected to safely recover at home and forgo a SNF stay.

---

[24] MLN Matters. Total knee arthroplasty (TKA) removal from the Medicare Inpatient Only (IPO) list and application of the 2-midnight rule. Last updated 24 January 2019. Available at: https://www.cms.gov/Outreach-and-Education/Medicare-Learning-Network-MLN/MLNMattersArticles/Downloads/SE19002.pdf

[25] Section II.E.1

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

[27] Centers for Medicare & Medicaid Services. Proposed Rule 42 CFR Part 510. Comprehensive Care for Joint Replacement Model, Three-Year Extension and Changes to Episode Definition and Pricing.

- If CJR participating hospitals move less costly patients to the outpatient setting, the remaining inpatient population would be costlier than the historical population used to calculate target prices (prior to Medicare coverage of OP TKA).

- Thus, shifting TKAs to the outpatient setting could make it harder for hospitals to reduce their payments below their target price and thus lower reconciliation payments for CJR participating hospitals.

■ Beginning in 2020, hospitals could be denied payment for inpatient TKAs that Recovery Audit Contractors (RACs) determine would have been appropriate for the outpatient setting.[28]

Which factors explain why mandatory CJR hospitals have a smaller OP TKA share than control group hospitals?

We examined whether the factors discussed above explain why the OP TKA share for mandatory CJR hospitals is 10 percentage points lower than the OP TKA share for the control group.

Patient complexity. Differences in patient complexity between mandatory CJR hospitals and control group hospitals, as measured by HCC, do not explain the difference in OP TKA shares. Exhibit 9 shows the hospital average OP TKA share by quartiles of the hospital average HCC of TKAs (inpatient and outpatient) in 2018. Since the OP TKA share for mandatory CJR hospitals is significantly lower than the control group OP TKA share at every HCC quartile, differences in patient complexity do not explain differences in OP TKA shares between mandatory CJR and control group hospitals. Moreover, while we hypothesized that the OP TKA share would decline as the HCC quartile increased, that was not the case for mandatory CJR hospitals or control group hospitals.

Exhibit 9:   Mandatory CJR hospitals had lower OP TKA shares across all quartiles of patient HCC score than the control group

Average HCC for TKA episodes, 2018	Mandatory CJR hospitals (n=270)		Control group hospitals (n=280)		
	Percent of hospitals	Average OP TKA share	Percent of hospitals	Average OP TKA share	p-value
<1.10	27.1%	21.1%	23.2%	35.1%	<0.01
1.10 - <1.23	22.1%	15.8%	29.1%	22.0%	<0.01
1.23 - <1.34	21.5%	19.7%	27.9%	34.0%	<0.01
≥1.34	29.3%	15.9%	19.9%	25.7%	<0.01

Source: CJR evaluation team analysis of Medicare claims and enrollment data for TKAs (inpatient and outpatient) initiated in 2017 and 2018.

Notes: Differences in averages were tested using t-tests. Results significant at the 99%, 95%, or 90% levels are indicated by red, orange, or yellow shaded cells, respectively.

---

[28]   The CY 2018 final rule prohibited Recovery Audit Contractor (RAC) patient status reviews for TKA procedures performed in the hospital inpatient setting for two years (CY 2018-2019) to allow hospitals time to adopt to the rule change without fear of payment denials.

This analysis includes hospitals with 10 or more inpatient TKAs in 2017 and is weighted by number of inpatient TKAs in 2017. Quartiles of HCC score were based on the 25th percentile, median, and 75th percentile.

HCC = hierarchical condition category, OP = outpatient, TKA = total knee arthroplasty.

Hospital length of stay prior to the OP TKA policy. Differences in hospital length of stay prior to the OP TKA policy do not explain the difference in OP TKA shares. Exhibit 10 shows the hospital OP TKA share by categories of the average hospital length of stay for TKAs in 2017. As we hypothesized, for both mandatory CJR hospitals and control group hospitals, hospitals with historically shorter lengths of stay performed a higher share of TKAs in the outpatient setting. For each length of stay category, however, the OP TKA share is significantly lower for mandatory CJR hospitals than for control group hospitals, indicating differences in length of stay do not explain differences in OP TKA shares between mandatory CJR and control group hospitals.

Exhibit 10:   Mandatory CJR hospitals have lower OP TKA shares than control group hospitals across all categories of length of stay for inpatient TKAs in 2017

Inpatient TKA LOS, 2017	Mandatory CJR hospitals (n=270)		Control group hospitals (n=280)		p-value
	Percent of hospitals	Mean OP TKA share, 2018	Percent of hospitals	Mean OP TKA share, 2018	
Less than 2 days	19.4%	43.9%	24.5%	45.9%	<0.01
2 days	65.1%	13.8%	67.8%	25.1%	<0.01
3 or more days	15.5%	3.9%	7.8%	11.0%	<0.01

Source: CJR evaluation team analysis of Medicare claims and enrollment data for TKAs (inpatient plus outpatient) initiated in 2017 and 2018.

Notes: Differences in averages were tested using t-tests. Results significant at the 99%, 95%, or 90% levels are indicated by red, orange, or yellow shaded cells, respectively.

LOS = length of stay, OP = outpatient, TKA = total knee arthroplasty.

Capacity to perform outpatient musculoskeletal procedures. We also hypothesized that hospitals that performed other musculoskeletal procedures in the outpatient setting prior to the OP TKA policy may transition TKAs more easily from the inpatient setting to the outpatient setting. We found, however, that nearly all mandatory CJR hospitals and control group hospitals were performing musculoskeletal procedures in the outpatient setting in 2017, so this did not help explain any differences.

Regional practice patterns. Regional practice patterns do not explain the difference in OP TKA shares. We hypothesized that the OP TKA shares of mandatory CJR and control group hospitals may reflect the OP TKA shares of hospitals in their hospital referral regions. We found that, on average, mandatory CJR hospitals had a lower OP TKA share than hospitals in their hospital referral region while control group hospitals had a higher OP TKA share than hospitals in their hospital referral region.

Multivariate analysis – patient, hospital, and regional differences. The OP TKA share for mandatory CJR hospitals remained 10 percentage points lower than the OP TKA share for control group hospitals, even after accounting for a variety of factors. We estimated risk-adjusted OP TKA shares adjusting for the same patient, hospital, and regional characteristics that we use for

our payment, service use, and quality impact estimates (including detailed HCC indicators and historical average hospital length of stay, see Appendix E, Exhibit E-15). Thus, participation in the CJR model, not patient complexity, length of stay, hospital or regional characteristics, appears to explain the difference in the share of TKAs performed in the outpatient setting between the mandatory CJR and control group hospitals. While we cannot say definitively that the CJR model is the cause of the difference, we have controlled for all other observable factors and ruled out several alternative explanations. Thus, the conservative conclusion is that the CJR model is affecting the adoption of TKA in the outpatient setting.

Impact of the TKA policy change on the CJR model evaluation

Under the CJR model, only inpatient LEJRs trigger CJR episodes. However, our analysis suggests that some inpatient TKA surgeries, which triggered CJR episodes, would have been outpatient procedures in the absence of CJR. Consequently, if we include only inpatient LEJR episodes in our analysis our impact estimates would be too high, that is, biased in favor of the CJR model, for two reasons: (1) they are based on an incorrect counterfactual, or assumption of what would have happened absent the model, and (2) the complexity of inpatient LEJR episodes may have increased less for CJR participant hospitals after the policy change than for control group hospitals. Therefore, for this annual report we also provide impact estimates based on episodes triggered by all LEJR procedures, including TKAs performed in the hospital outpatient setting.

Incorrect counterfactual. An impact estimate that includes only episodes triggered by an inpatient LEJR is based on the wrong counterfactual for CJR inpatient TKA episodes. As detailed above, the CJR episodes include patients who likely would have had outpatient TKAs based on statistical modeling, and thus not triggered CJR episodes, in the absence of the CJR model. These episodes incurred higher episode payments, both because payment to the hospital for the inpatient procedure is higher than the payment for the outpatient procedure and because some of these included SNF payments, which would not be covered following an outpatient TKA. An impact analysis including only inpatient LEJRs would incorrectly compare payments and use for these TKAs with payments and use for inpatient TKAs from the control group, and thus would not detect these payment and use increases.

Differential change in patient mix. Patients who could appropriately receive TKAs as outpatients are expected to be less complex and thus require fewer resources than patients who require TKAs to be performed in the inpatient setting. As control group hospitals perform more TKAs in the outpatient setting than mandatory CJR hospitals, the inpatient patient mix of control group hospitals would be expected to become relatively more complex than the inpatient patient mix of mandatory CJR hospitals. As discussed in Section II.B.6.b, we did find a relative increase in patient complexity for inpatients at control group hospitals. That differential change in patient mix could bias the impact estimates that are based on only inpatient LEJR episodes to indicate too large of a reduction in payments and use, despite our rigorous risk adjustment methodology.

d. Conclusion

The CJR model provides an additional financial incentive to perform TKAs in the inpatient setting. Indeed, mandatory CJR hospitals performed a smaller proportion of TKAs in the outpatient setting than control group hospitals. Evidence indicates that mandatory CJR hospitals' lower uptake of OP TKA is due to the CJR model. The impact estimates based only on inpatient LEJR episodes use an incorrect counterfactual for 10% of TKAs in PY3. Therefore, the impact estimates based only on inpatient LEJR episodes are biased and overestimate the impact of the CJR model on reductions in payments and service use. In this annual report we also provide impact estimates based on all LEJRs, including OP TKAs. The impact estimates based on all LEJRs, including OP TKAs, account for the impact of the CJR model on the decision about whether the TKA was performed in the inpatient or the outpatient setting. Because they also include the impact on outpatient TKAs that are not subject to CJR treatment, the impact estimates based on all LEJRs, including OP TKAs, likely underestimate the impact of the CJR model on payments and service use. However, the impact estimates based on all LEJRs, including OP TKAs, are likely closer to the actual impact of the CJR model because they consider the differential outpatient TKA rate, which our analyses suggest was due to the CJR model.

B. Impact of the Model

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

Changes to the CJR model that took effect in the third performance year introduce additional opportunities and challenges in understanding the impact of the CJR model. First, CMS removed TKA from the inpatient only list, allowing Medicare coverage for TKAs provided in the hospital outpatient setting. Evidence suggests that the CJR model influences the choice of inpatient or outpatient setting, which would bias impact estimates that are based only on inpatient LEJR episodes.[29] Therefore, in this section, we provide impact estimates based on all LEJRs (IP+OP) to account for the impact of the CJR model on the inpatient or outpatient decision (see definitions in the box below and Section II.A.2 for additional information). Impact estimates based on all LEJRs (IP+OP) provide an underestimate of the impact for two reasons. First, they include outpatient TKAs which are not episodes under the CJR model, and therefore, likely were not subject to CJR treatment. Second, average payments for OP TKAs in 2018 were higher in CJR participant hospitals than for the control group hospitals. Because of that, we also report impact estimates based on only inpatient LEJR episodes. These estimates overestimate the impact of the CJR model because they do not consider the lower proportion of outpatient TKA surgeries performed in CJR participant hospitals. Because our analyses suggest that the lower outpatient TKA rate in mandatory CJR hospitals was due to the CJR model, the estimate based on all LEJRs (IP+OP) is closer to the truth as it considers the differential outpatient TKA rate.

---

[29] Centers for Medicare & Medicaid Services. Proposed Rule 42 CFR Part 510. Comprehensive Care for Joint Replacement Model, Three-Year Extension and Changes to Episode Definition and Pricing.

An inpatient LEJR episode begins with an inpatient anchor hospitalization that meets CJR episode eligibility requirements and ends 90 days after discharge. Inpatient LEJR episodes are CJR episodes under the CJR model.

Outpatient TKA begins with an outpatient anchor procedure that meets CJR episode eligibility requirements and ends 90 days after the date of service for the anchor procedure. Outpatient TKAs are not CJR episodes under the CJR model.

All LEJRs (IP+OP) include inpatient LEJR episodes and outpatient TKAs.

Impact estimates based on
- All LEJRs (IP+OP) underestimate the impact of the CJR model on average episode payments.
- Inpatient LEJR episodes overestimate the impact of the CJR model on average episode payments.

Second, as discussed in Sections I.A and II.A.1, the number of mandatory MSAs was scaled back to the 34 MSAs with the highest historical payments. While this evaluation report focuses on the MSAs and 378 hospitals that were always required to participate, in this section, we also estimate the impact of the CJR model on average episode payments for the 758 hospitals in the original 67 CJR MSAs. This analysis provides insight into the impact of the CJR model on hospitals included in the original design of the model, spanning the two years in which they participated and the one year following. In the next annual report, we will estimate the impact of the CJR model on average episode payments separately for hospitals in voluntary MSAs that opted into the CJR model and for hospitals that did not opt into the CJR model.

a. Key Findings

■ For mandatory CJR hospitals, average payments (standardized allowed amounts) for *all LEJRs (IP+OP)* decreased by $1,378 more than for control group hospitals during the first three performance years of the CJR model. This equates to a 4.7% decrease from the baseline. Average payments for *inpatient LEJR episodes* decreased by $1,540 more than for control group hospitals during the first three performance years. This equates to a 5.3% decrease from the baseline. The estimate for *all LEJRs (IP+OP)* is an underestimate, while the estimate for *inpatient LEJR episodes* is an overestimate.

■ For hospitals in the original CJR MSAs, average payments for *all LEJRs (IP+OP)* decreased by $866 more than for control group hospitals during the first three performance years of the CJR model. This equates to a 3.2% decrease from the baseline. Average payments for *inpatient LEJR episodes* decreased by $991 more than for control group hospitals during the first three performance years. This equates to a 3.7% decrease from the baseline. The estimate for *all LEJR episodes* is an underestimate, while the estimate for *inpatient LEJR episodes* is an overestimate.

■ For mandatory CJR hospitals and hospitals in the original CJR MSAs, the reduction in average payments for all LEJRs (IP+OP) for PY3 was smaller than the reduction in PY2 due to the impact of the CJR model on the inpatient or outpatient decision.

b. Methods

The analysis used 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 2018, or the first three performance years of the CJR model) for all LEJRs (IP+OP) and inpatient LEJR episodes initiated at mandatory CJR hospitals and hospitals in the original CJR MSAs relative to those initiated at control group hospitals. We used standardized payments to ensure that observed payment differences reflect actual differences in billed services rather than Medicare payment policies. We used allowed amounts to eliminate variation in payments due to whether beneficiaries have met their deductible when they had the LEJR surgery. We used the DiD method because it controls for trends that may affect both CJR and control group hospitals. In addition, we risk-adjusted estimates for beneficiary, market, and hospital characteristics that can vary over time and between the CJR and control group. We assessed the impact of the CJR model on average episode payments for mandatory CJR hospitals and hospitals in the original CJR MSAs.[30] The control group MSAs were weighted to be representative of the distribution of the CJR MSAs of interest.

---

[30] Of the 236,153 inpatient LEJR episodes generated by the 758 hospitals in the original CJR MSAs, 89% were generated by hospitals while they were participating in the CJR model; 11% were generated by hospitals after they stopped participating in the model. For all LEJRs (IP+OP), 88% were generated by hospitals while they were participating in the CJR model; 12% were generated by hospitals after they stopped participating in the CJR model.

The percent decrease in payments represents the percent change from the CJR baseline that is due to the CJR model. It is calculated by dividing the DiD estimate by the CJR baseline average.

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

While the DiD method is intended to isolate the impact of the CJR model, it does not fully control for all external influences. Specifically, it does not fully control for the unequal contribution of episodes from former BPCI participants across the CJR and control group samples. Hospitals in CJR MSAs dropped out of BPCI earlier and at a higher rate than hospitals in control group MSAs. That led to a larger contribution of intervention episodes by former BPCI hospitals in the CJR group compared to the control group.[31] Because former BPCI hospitals have lower average episode payments during the CJR performance period, their higher contribution to the CJR group may exaggerate the reduction in average episode payments for the CJR group during the intervention. Our robustness analyses indicate that the larger contribution of episodes by former BPCI participants may overestimate the reductions in average episode payment due to CJR by roughly $200 per episode.

Additional details about the methodology, including this limitation and the robustness analyses, is in Appendix E (Section III.D.1).

c.   Results

During each of the first three performance years, the CJR model resulted in relative reductions in average episode payments. This holds for the two samples studied—all LEJRs (IP+OP) and inpatient LEJR episodes—and for the two hospital groups studied—mandatory CJR hospitals and hospitals in the original CJR MSAs (Exhibit 11). The analysis of average episode payments does not incorporate reconciliation payments made to CJR participant hospitals; therefore, the results do not represent savings to the Medicare program. An analysis of Medicare savings is in Section II.B.2.

---

[31]   Episodes contributed by hospitals participating in BPCI are excluded from the analysis. Hospitals participating in BPCI in control group MSAs were less likely to drop out of BPCI and so more of their episodes are excluded from the CJR analysis.

Exhibit 11: **Impact of the CJR model on average episode payments, PY1-PY3**

	Mandatory CJR hospitals	Hospitals in the original CJR MSAs
All LEJRs (IP+OP) (underestimate)	-$1,378 (-4.7%)	-$866 (-3.2%)
Inpatient LEJR episodes (overestimate)	-$1,540 (-5.3%)	-$991 (3.7%)

Source: CJR evaluation team analysis of Medicare claims and enrollment data for LEJRs initiated in 2012 through 2014 that ended between April 2012 and March 2015 (baseline) and LEJRs initiated during or after April 2016 that ended by December 2018 (intervention).

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

All LEJRs (IP+OP) include inpatient LEJR episodes and outpatient TKAs. An inpatient LEJR episode begins with an inpatient anchor hospitalization that meets CJR episode eligibility requirements and ends 90 days after discharge. Inpatient LEJR episodes are CJR episodes under the CJR model. An outpatient TKA begins with an outpatient anchor procedure that meets CJR episode eligibility requirements and ends 90 days after the date of service for the anchor procedure. Outpatient TKAs are not CJR episodes under the CJR model.

Impact estimates based on all LEJRs (IP+OP) underestimate the impact of the CJR model on average episode payments.

Impact estimates based on inpatient LEJR episodes overestimate the impact of the CJR model on average episode payments.

IP = inpatient, LEJR = lower extremity joint replacement, MSA = metropolitan statistical area, OP = outpatient, PY = performance year, TKA = total knee arthroplasty.

All LEJRs (IP+OP)

Mandatory CJR hospitals

While average payments for all LEJRs (IP+OP) declined for both mandatory CJR hospitals and control group hospitals during the first three performance years, payments declined more for all LEJRs (IP+OP) initiated at mandatory CJR hospitals (Exhibit 12). Average payments decreased by $1,378 more for all LEJRs (IP+OP) initiated at mandatory CJR hospitals than for all LEJRs (IP+OP) initiated at control group hospitals from the baseline to the intervention period ($p<0.01$, Exhibit 11). This relative reduction equates to a 4.7% decrease from the baseline average payments for all LEJRs (IP+OP) at mandatory CJR hospitals.

Exhibit 12: Average payments declined more for <u>all LEJRs (IP+OP)</u> at mandatory CJR hospitals than for all LEJRs (IP+OP) at control group hospitals in PY1-3*

Source: CJR evaluation team analysis of Medicare claims and enrollment data for all LEJRs (IP+OP) initiated on or after January 2012 that ended by December 2018.

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

All LEJRs (IP+OP) include inpatient LEJR episodes and outpatient TKAs. An inpatient LEJR episode begins with an inpatient anchor hospitalization that meets CJR episode eligibility requirements and ends 90 days after discharge. Inpatient LEJR episodes are CJR episodes under the CJR model. An outpatient TKA begins with an outpatient anchor procedure that meets CJR episode eligibility requirements and ends 90 days after the date of service for the anchor procedure. Outpatient TKAs are not CJR episodes under the CJR model.

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

IP = inpatient, LEJR = lower extremity joint replacement, OP = outpatient, PY= performance year, TKA = total knee arthroplasty.

For mandatory CJR hospitals, average payments for all LEJRs (IP+OP) decreased in each of the three performance years, however, the decrease in PY3 was smaller than the decrease in PY2 (p<0.05). Average payments decreased by $1,401 more for all LEJRs (IP+OP) at mandatory CJR hospitals than for all LEJRs (IP+OP) at control group hospitals in the first performance year (p<0.01), by $1,605 more in the second performance year (p<0.01), and by $1,147 more in the third performance year (p<0.01) (Exhibit 13).

Exhibit 13: **Average payments for <u>all LEJRs (IP+OP)</u> at mandatory CJR hospitals decreased in all three performance years, but the decrease in PY3 was smaller than the decrease in PY2***

*Impact estimates based on all LEJRs (IP+OP), which underestimate the impact of the CJR model

Source: CJR evaluation team analysis of Medicare claims and enrollment data for all LEJRs (IP+OP) initiated in 2012 through 2014 that ended between April 2012 and March 2015 (baseline) and all LEJRs (IP+OP) initiated during or after April 2016 that ended by December 2018 (intervention).

Notes: The estimates in this exhibit are the result of a DiD model. DiD estimates that are significant at the <mark style="background:#8B1A2B;color:white">99%</mark>, <mark style="background:orange">95%</mark>, or <mark style="background:#F5C518">90%</mark> significance level are indicated by red, orange, or yellow shaded circles, respectively. Confidence intervals are set at the 90% level.

All LEJRs (IP+OP) include inpatient LEJR episodes and outpatient TKAs. An inpatient LEJR episode begins with an inpatient anchor hospitalization that meets CJR episode eligibility requirements and ends 90 days after discharge. Inpatient LEJR episodes are CJR episodes under the CJR model. An outpatient TKA begins with an outpatient anchor procedure that meets CJR episode eligibility requirements and ends 90 days after the date of service for the anchor procedure. Outpatient TKAs are not CJR episodes under the CJR model.

Impact estimates based on all LEJRs (IP+OP) underestimate the impact of the CJR model on average episode payments.

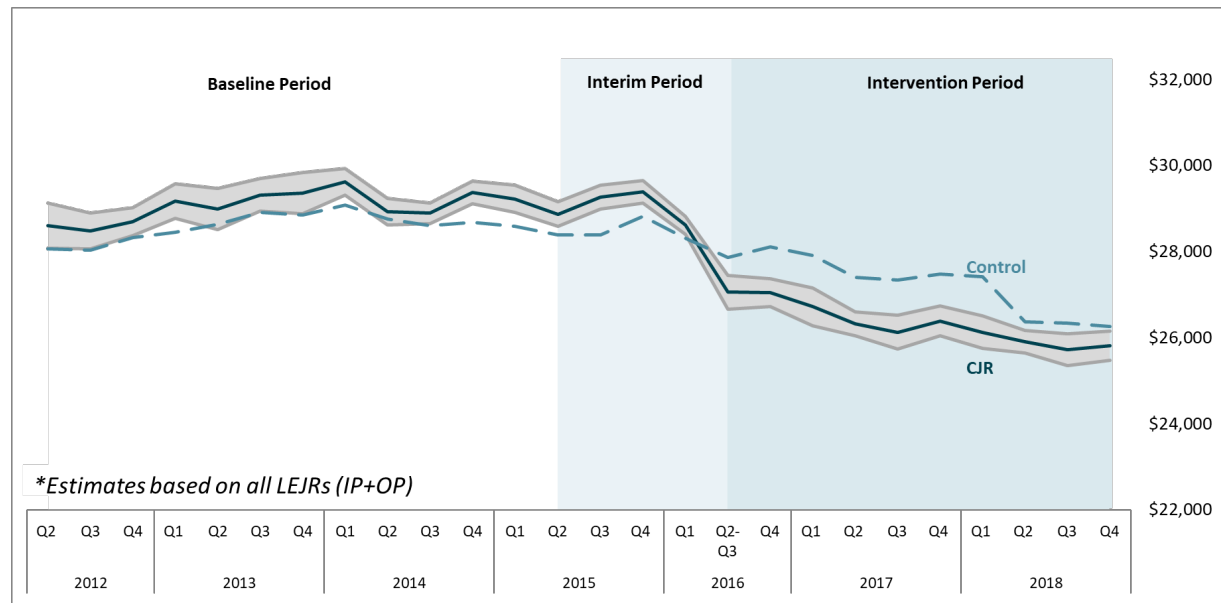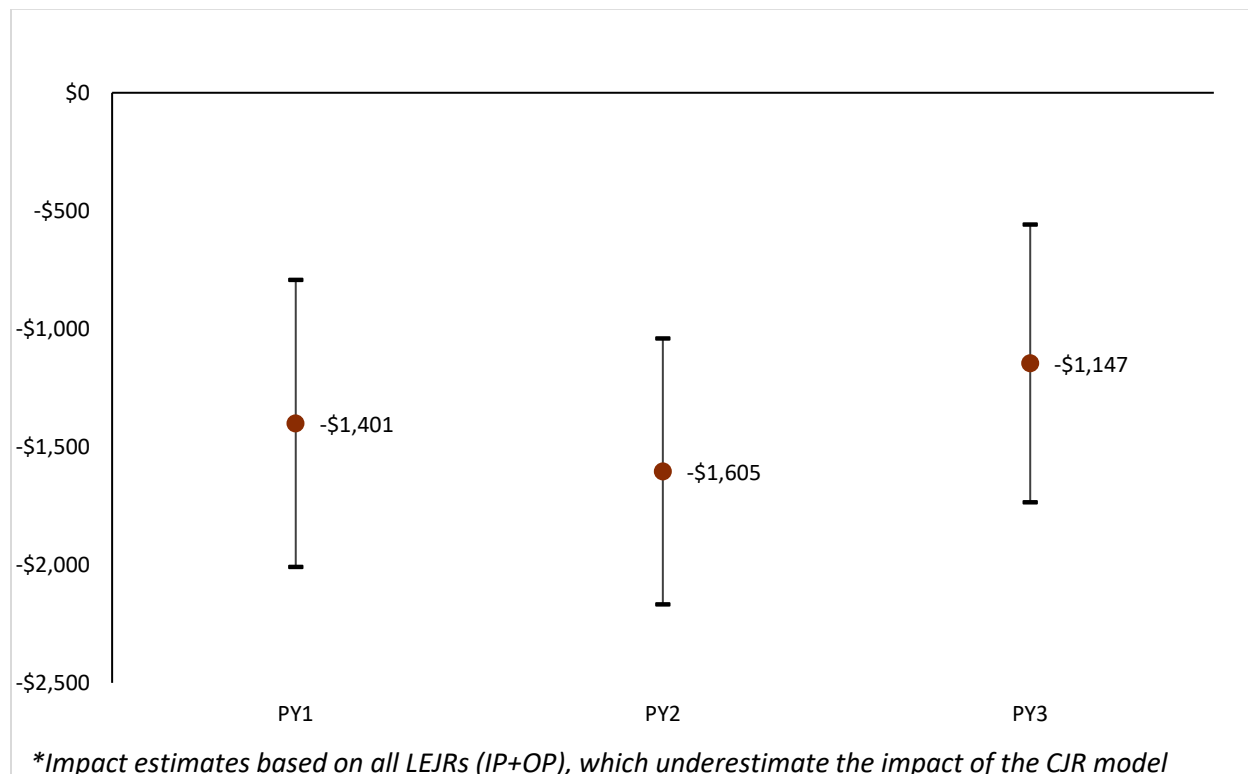DiD = difference-in-differences, IP = inpatient, LEJR = lower extremity joint replacement, OP = outpatient, PY = performance year, TKA = total knee arthroplasty.

Hospitals in the original CJR MSAs

Average payments for all LEJRs (IP+OP) declined for both hospitals in the original CJR MSAs and control group hospitals during the first three performance years, but payments declined more for all LEJRs (IP+OP) initiated at hospitals in the original CJR MSAs (Exhibit 14). Average payments decreased by $866 more for all LEJRs (IP+OP) initiated at hospitals in the original CJR MSAs than for all LEJRs (IP+OP) initiated at control group hospitals from the baseline to the intervention period (p<0.01, Exhibit 11). This relative reduction equates to a 3.2% decrease from the baseline in average payments for all LEJRs (IP+OP) at hospitals in the original CJR MSAs.

For hospitals in the original CJR MSAs, average payments for all LEJRs (IP+OP) decreased in each of the three performance years, however, the point estimate in PY3 was significantly smaller than the point estimate in PY1 (p<0.10) and in PY2 (p<0.01). Average payments decreased by $884 more for all LEJRs (IP+OP) at hospitals in the original CJR MSAs than for all LEJRs (IP+OP) at control group hospitals in the first performance year (p<0.01), by $1,065 more in the second performance year (p<0.01), and by $637 more in the third performance year (p<0.05) (Exhibit 15).

The reduction in average episode payments was larger for mandatory CJR hospitals than for hospitals in the original CJR MSAs. By design, mandatory CJR hospitals had higher historical average episode payments than hospitals in the original CJR MSAs suggesting that mandatory CJR hospitals had greater opportunity to reduce average episode payments. Additionally, hospitals in the mandatory MSAs continue to face the CJR incentives. For hospitals in the original CJR MSAs, 12% of episodes were generated by hospitals after they were no longer participating in the CJR model.

Exhibit 14:   Average payments declined more for all LEJRs (IP+OP) at hospitals in the original CJR MSAs than for all LEJRs (IP+OP) at control group hospitals in PY1-3*

Estimates based on all LEJRs (IP+OP)

Source: CJR evaluation team analysis of Medicare claims and enrollment data for all LEJRs (IP+OP) initiated on or after January 2012 that ended by December 2018.

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

All LEJRs (IP+OP) include inpatient LEJR episodes and outpatient TKAs. An inpatient LEJR episode begins with an inpatient anchor hospitalization that meets CJR episode eligibility requirements and ends 90 days after discharge. Inpatient LEJR episodes are CJR episodes under the CJR model. An outpatient TKA begins with an outpatient anchor procedure that meets CJR episode eligibility requirements and ends 90 days after the date of service for the anchor procedure. Outpatient TKAs are not CJR episodes under the CJR model.

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

IP = inpatient, LEJR = lower extremity joint replacement, MSA = metropolitan statistical area, OP = outpatient, PY = performance year, TKA = total knee arthroplasty.

Exhibit 15: **Average payments for <u>all LEJRs (IP+OP)</u> at hospitals in the original CJR MSAs decreased in all three performance years, but the decrease in PY3 was smaller than the decreases in PY1 and PY2***

*Impact estimates based on all LEJRs (IP+OP), which underestimate the impact of the CJR model

Source: CJR evaluation team analysis of Medicare claims and enrollment data for all LEJRs (IP+OP) initiated in 2012 through 2014 that ended between April 2012 and March 2015 (baseline) and all LEJRs (IP+OP) initiated during or after April 2016 that ended by December 2018 (intervention).

Notes: The estimates in this exhibit are the result of a DiD model. DiD estimates that are significant at the 99%, 95%, or 90% significance level are indicated by red, orange, or yellow shaded circles, respectively. Confidence intervals are set at the 90% level.

All LEJRs (IP+OP) include inpatient LEJR episodes and outpatient TKAs. An inpatient LEJR episode begins with an inpatient anchor hospitalization that meets CJR episode eligibility requirements and ends 90 days after discharge. Inpatient LEJR episodes are CJR episodes under the CJR model. An outpatient TKA begins with an outpatient anchor procedure that meets CJR episode eligibility requirements and ends 90 days after the date of service for the anchor procedure. Outpatient TKAs are not CJR episodes under the CJR model.

Impact estimates based on all LEJRs (IP+OP) underestimate the impact of the CJR model on average episode payments.

DiD = difference-in-differences, IP = inpatient, LEJR = lower extremity joint replacement, MSA = metropolitan statistical area, OP = outpatient, PY = performance year, TKA = total knee arthroplasty.

Inpatient LEJR episodes

Mandatory CJR hospitals

While average payments for inpatient LEJR episodes declined for both mandatory CJR hospitals and control group hospitals during the first three performance years, payments declined more for inpatient LEJR episodes initiated at mandatory CJR hospitals (Exhibit 16). Average payments decreased by $1,540 more for inpatient LEJR episodes initiated at mandatory CJR hospitals than for inpatient LEJR episodes initiated at control group hospitals from the baseline to the intervention period (p<0.01, Exhibit 11). This relative reduction equates to a 5.3% decrease from the baseline in average payments for inpatient LEJR episodes at mandatory CJR hospitals.

Exhibit 16: **Average payments declined more for <u>inpatient LEJR episodes</u> at mandatory CJR hospitals than for inpatient LEJR episodes at control group hospitals in PY1-3***

Source: CJR evaluation team analysis of Medicare claims and enrollment data for inpatient LEJR episodes initiated on or after January 2012 that ended by December 2018.
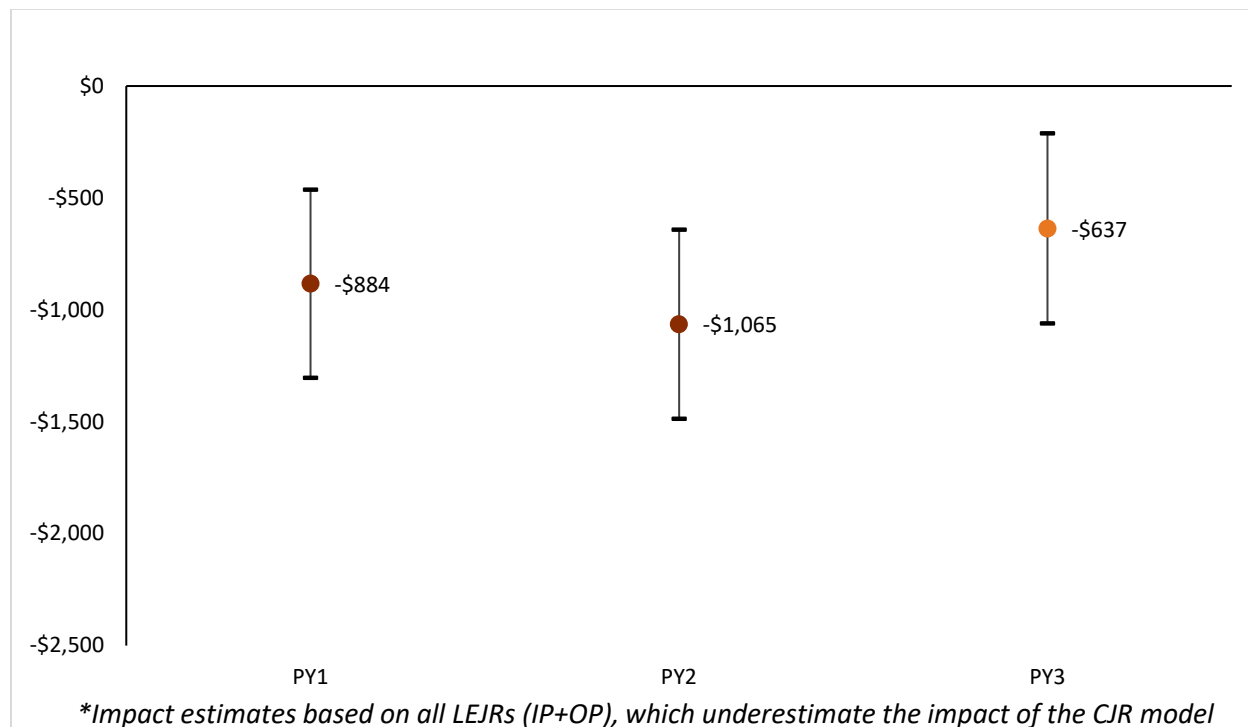
Notes: Inpatient LEJR episodes that ended between April 1, 2015 and March 31, 2016 (the interim period) were excluded from our baseline because the CJR model was announced in July 2015 and hospitals may have been preparing for their future participation in the CJR model during that time.

An inpatient LEJR episode begins with an inpatient anchor hospitalization that meets CJR episode eligibility requirements and ends 90 days after discharge. Inpatient LEJR episodes are CJR episodes under the CJR model.

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

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

For mandatory CJR hospitals, average payments for inpatient LEJR episodes decreased in each of the three performance years, and the reduction was stable over time. Average payments decreased by $1,402 more for inpatient LEJR episodes at mandatory CJR hospitals than for inpatient LEJR episodes at control group hospitals in the first performance year (p<0.01), by $1,605 more in the second performance year (p<0.01), and by $1,531 more in the third performance year (p<0.01, Exhibit 17). The point estimates for each performance year were not statistically different from one another (p=0.59). This differs from what we observed for the all LEJRs (IP+OP) sample, which showed a statistically significantly smaller point estimate in PY3 ($1,147) compared to PY2 ($1,605, p<0.05). The all LEJRs (IP+OP) result considers the impact of the CJR model on the inpatient or outpatient decision.

Exhibit 17: **Average payments for <u>inpatient LEJR episodes</u> at mandatory CJR hospitals decreased in PY1, PY2, and PY3***

*Impact estimates based on inpatient LEJRs, which overestimate the impact of the CJR model

Source: CJR evaluation team analysis of Medicare claims and enrollment data for inpatient LEJR episodes initiated in 2012 through 2014 that ended between April 2012 and March 2015 (baseline) and inpatient LEJR episodes initiated during or after April 2016 that ended by December 2018 (intervention).

Notes: The estimates in this exhibit are the result of a DiD model. DiD estimates that are significant at the 99%, 95%, or 90% significance level are indicated by red, orange, or yellow shaded circles, respectively. Confidence intervals are set at the 90% level.

An inpatient LEJR episode begins with an inpatient anchor hospitalization that meets CJR episode eligibility requirements and ends 90 days after discharge. Inpatient LEJR episodes are CJR episodes under the CJR model.

Impact estimates based on inpatient LEJR episodes overestimate the impact of the CJR model on average episode payments.

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

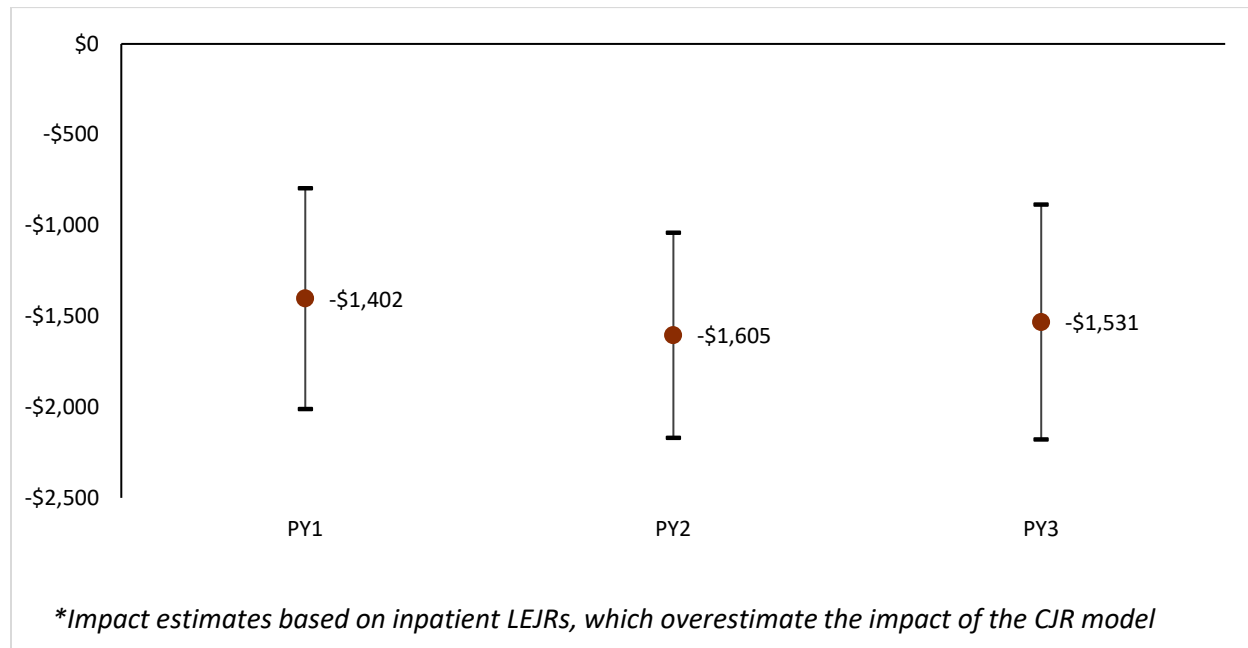## *Hospitals in the original CJR MSAs*

While average payments for inpatient LEJR episodes declined for both hospitals in the original CJR MSAs and in the control group during the first three performance years, payments declined more for inpatient LEJR episodes initiated at hospitals in the original CJR MSAs (Exhibit 18). Average payments decreased by $991 more for inpatient LEJR episodes initiated at hospitals in the original CJR MSAs than for inpatient LEJR episodes initiated at control group hospitals from the baseline to the intervention period (p<0.01, Exhibit 11). This relative reduction equates to a 3.7% decrease from the baseline in average payments for inpatient LEJR episodes at hospitals in the original CJR MSAs.

For hospitals in the original CJR MSAs, average payments for inpatient LEJR episodes decreased in each of the three performance years, and the reduction was stable over time. Average payments decreased by $884 more for inpatient LEJR episodes at hospitals in the original CJR MSAs than for inpatient LEJR episodes at control group hospitals in the first performance year (p<0.01), by $1,063 more in the second performance year (p<0.01), and by $946 more in the third performance

year (p<0.01) (Exhibit 19). The point estimates for each performance year were not statistically different from one another (p=0.28).

Exhibit 18:   Average payments declined more for <u>inpatient LEJR episodes</u> at hospitals in the original CJR MSAs than for inpatient LEJR episodes at control group hospitals in PY1-3*

Source: CJR evaluation team analysis of Medicare claims and enrollment data for inpatient LEJR episodes initiated on or after January 2012 that ended by December 2018.

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

An inpatient LEJR episode begins with an inpatient anchor hospitalization that meets CJR episode eligibility requirements and ends 90 days after discharge. Inpatient LEJR episodes are CJR episodes under the CJR model.

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

LEJR = lower extremity joint replacement, MSA = metropolitan statistical area, PY = performance year.

Exhibit 19: **Average payments for <u>inpatient LEJR episodes</u> at hospitals in the original CJR MSAs decreased in PY1, PY2, and PY3***

*Impact estimates based on inpatient LEJRs, which overestimate the impact of the CJR model

Source: CJR evaluation team analysis of Medicare claims and enrollment data for inpatient LEJR episodes initiated in 2012 through 2014 that ended between April 2012 and March 2015 (baseline) and inpatient LEJR episodes initiated during or after April 2016 that ended by December 2018 (intervention).

Notes: The estimates in this exhibit are the result of a DiD model. DiD estimates that are significant at the 99%, 95%, or 90% significance level are indicated by red, orange, or yellow shaded circles, respectively. Confidence intervals are set at the 90% level.

An inpatient LEJR episode begins with an inpatient anchor hospitalization that meets CJR episode eligibility requirements and ends 90 days after discharge. Inpatient LEJR episodes are CJR episodes under the CJR model.

Impact estimates based on inpatient LEJR episodes overestimate the impact of the CJR model on average episode payments.

DiD = difference-in-differences, LEJR = lower extremity joint replacement, MSA = metropolitan statistical area, PY = performance year.
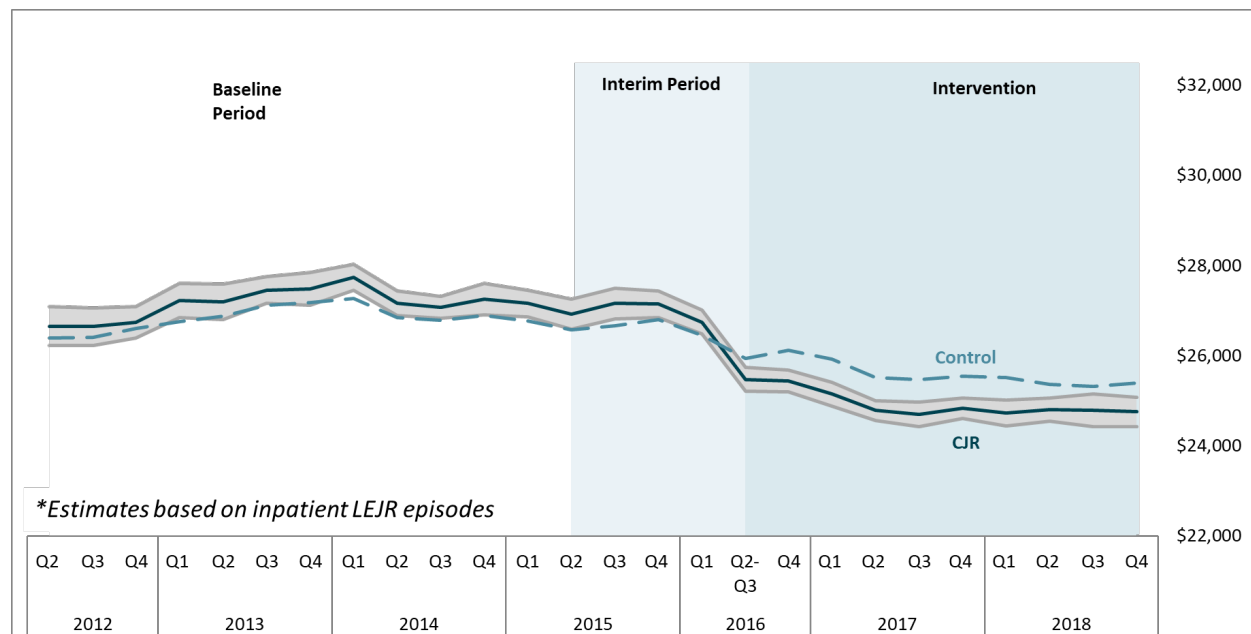
For both mandatory CJR hospitals and hospitals in the original CJR MSAs, the reductions in average payments for inpatient LEJR episodes were generally stable from PY2 to PY3 because they do not consider the impact of the CJR model on the inpatient or outpatient decision in PY3.

d. Conclusion

Over the first three performance years, the CJR model resulted in relative reductions in average episode payments for both samples studied—all LEJRs (IP+OP) and inpatient LEJR episodes—and for both hospital groups studied—mandatory CJR hospitals and hospitals in the original CJR MSAs. There are differences in the results for the two samples and for the two hospital groups, however.

First, the relative reductions in average episode payments were smaller in magnitude for all LEJRs (IP+OP) than for inpatient LEJR episodes because the estimates based on all LEJRs (IP+OP) consider the impact of the CJR model on the inpatient or outpatient decision. Second, the relative reductions in average episode payments were larger in magnitude for mandatory CJR hospitals than for hospitals in the original 67 MSAs because mandatory CJR hospitals historically had higher episode payments and more opportunity to reduce episode payments under the CJR model. Further, hospitals in the original 67 MSAs included hospitals in voluntary MSAs that ended their participation in the CJR model in PY3.

Differences between the two samples were also evident when looking specifically at estimates for PY3. For all LEJRs (IP+OP), the decrease in average episode payments were smaller for PY3 than PY2 for both mandatory CJR hospitals and hospitals in the original 67 MSAs. That is because PY3 coincided with the removal of TKAs from the inpatient only list and the estimates for all LEJRs (IP+OP) reflect the lower proportion of outpatient TKAs at CJR participant hospitals in PY3 compared to control group hospitals. The estimates for all LEJRs (IP+OP) also likely underestimate the impact of the CJR model in PY3 because they also include the impact on outpatient TKAs that likely were not subject to CJR treatment. For inpatient LEJRs, the decrease in average payments remained stable from PY2 to PY3 for both mandatory CJR hospitals and hospitals in the original 67 MSAs. The estimates for inpatient LEJRs are likely overestimating the impact of the CJR model in PY3 since they do not consider the lower proportion of outpatient TKAs at CJR participant hospitals compared to control group hospitals. The true impact for PY3 likely falls between these two estimates.

It is likely that the decrease in average episode payments will be smaller for the fourth and fifth performance years as more time passes under the OP TKA payment policy and as Medicare coverage of THA expands to include the hospital outpatient department and TKA is covered in ambulatory surgery centers (beginning with PY5). In part in response to these changes, CMS published a proposed rule in February 2020 that would extend the CJR model for an additional three years and expand the CJR episode definition during the extension to include OP TKAs and OP THAs.[32]

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

Medicare savings from the CJR model reflects the change in average episode payments as well as the reconciliation payments made to or received from hospitals under the model. As presented in Section II.B.1, mandatory CJR hospitals reduced average standardized episode payments for LEJR episodes during the first three performance years. This section presents estimated Medicare savings based on the change in average non-standardized paid amounts and reconciliation payments made to or received from mandatory CJR participant hospitals.

---

[32]   Centers for Medicare & Medicaid Services. Proposed Rule 42 CFR Part 510. Comprehensive Care for Joint Replacement Model, Three-Year Extension and Changes to Episode Definition and Pricing.

> ### a. Key Findings
>
> - During the first three performance years, estimated Medicare savings for mandatory CJR hospitals, based on *all LEJRs (IP+OP)*, was $61.6 million, ranging from savings of $2.3 million to $120.9 million. This is an underestimate, and constitutes a lower bound.
>
>   - The estimate of savings was $536 per LEJR (IP+OP), ranging from savings of $20 to $1,052.
>
> - During the first three performance years, estimated Medicare savings for *inpatient LEJR episodes* at mandatory CJR hospitals was $76.3 million, ranging from savings of $16.0 million to $136.7 million. This is an overestimate, and constitutes an upper bound.
>
>   - The estimate of savings was $685 per episode, ranging from savings of $143 to $1,227.

b. Methods

We provide Medicare savings estimates based on all LEJRs, including OP TKAs, to account for the impact of the CJR model on the inpatient or outpatient decision after Medicare began covering TKAs provided in a hospital outpatient setting at the start of PY3 (See Section II.A.2). Medicare savings estimates based on all LEJRs (IP+OP) likely underestimate the true impact of the CJR model for two reasons. First, they include outpatient TKAs, which are not episodes under the CJR model, and therefore likely were not subject to CJR treatment. Second, average payments for OP TKAs in 2018 were higher in mandatory CJR hospitals than for the control group hospitals. We also report Medicare savings estimates based on only inpatient LEJR episodes. These estimates likely overestimate the true impact of the CJR model because they assume that the lower outpatient TKA rate in mandatory CJR hospitals was not due to the model. Because our analyses indicate that the lower outpatient TKA rate in mandatory CJR hospitals was due to the CJR model, the estimate based on all LEJRs (IP+OP) is closer to the truth.

Additionally, as discussed in Sections I.A and II.A.1, the number of mandatory MSAs was scaled back to the 34 MSAs with the highest historical average episode payments. This section focuses on those MSAs and the 378 hospitals that were always required to participate. Our calculations of Medicare savings are for the mandatory CJR hospitals, and therefore, do not represent Medicare savings attributable to all hospitals that ever participated in the CJR model. Future reports will study the effect of the CJR model on the hospitals in the voluntary MSAs.

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

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

The change in non-standardized paid amounts is calculated using our difference-in-differences methodology, explained fully in the Methods appendix (Appendix E, Section III.D.1).

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

Reconciliation payments are the payments made to CJR participant hospitals by Medicare for meeting cost and quality targets and repayments from CJR participant hospitals to Medicare for failing to meet cost and quality targets.[33] See Appendix E (Section V.B) for additional details about these methods and limitations of this methodology.

c.   Results

Savings estimates based on all LEJRs (IP+OP)

For all LEJRs (IP+OP) at mandatory CJR hospitals, the CJR model reduced Medicare program spending by an estimated $61.6 million during the first three performance years (Exhibit 20). After accounting for the uncertainty in the reduction in per-LEJR (IP+OP) payments, the estimated savings due to the CJR model ranged from savings of $2.3 million to $120.9 million. Medicare savings is based on an estimated reduction in per-LEJR (IP+OP) non-standardized paid amounts of $1,323 for the 114,907 LEJRs that were performed at CJR-participant hospitals on either an inpatient or outpatient basis, or $152.1 million. From this, we subtracted reconciliation payments. To obtain reconciliation payments per LEJR (IP+OP), we divided total reconciliation payments of $90.5 million by the 114,907 LEJRs (IP+OP), resulting in reconciliation payments per LEJR (IP+OP) of $787. Subtracting reconciliation payments from reductions in payments results in an estimated savings of $536 per LEJR (IP+OP), ranging from savings of $20 to $1,052, or in aggregate, savings of $61.6 million ranging from savings of $2.3 million to $120.9 million. The point estimate equates to savings of 2.0% of the baseline.

---

[33]   Reconciliation payments can be positive or negative, depending on if the payment is from Medicare to a participant or a repayment from the participant to Medicare. In the CJR model rule these payments are often referred to by the technical term "net payment reconciliation amounts" or "NPRA."

Exhibit 20: The CJR model resulted in savings to Medicare for mandatory CJR hospitals in PY1-3 (underestimated)

Component	Per LEJR (IP+OP) Average		Number of LEJRs (IP+OP)	Total	
	Estimate	Range		Estimate	Range
Reduction in non-standardized paid amounts for all LEJRs (IP+OP)	$1,323	$807 to $1,840	114,907	$152,065,295	$92,732,623 to $211,397,968
Reconciliation payments	$787	--	114,907	$90,476,027	--
Medicare savings	$536	$20 to $1,052	114,907	$61,589,268	$2,256,596 to $120,921,941

Source: CJR evaluation team analysis of Medicare claims and enrollment data for episodes initiated in 2012 through 2014 that ended between April 2012 and March 2015 (baseline) and LEJRs initiated during or after April 2016 that ended by December 2018 (intervention) and CJR payment contractor data for CJR participant hospitals in PY1-3.

Notes: Reductions in non-standardized paid amounts are based on a weighted average of performance year estimates from a DiD model of per-LEJR standardized paid amounts that have been multiplied by negative one and converted to non-standardized amounts. Ranges are based on 90% confidence intervals. Results are based on all LEJRs (IP+OP) and are underestimated.

DiD = difference-in-differences, IP = inpatient, LEJR = lower extremity joint replacement, OP = outpatient, PY=performance year.

We also evaluated Medicare savings separately for each performance year to better understand how model features that changed across performance years and the removal of TKA from the inpatient only list affected savings. The stop-gain limit increased from 5% in PY1 to 10% in PY2, allowing hospitals to receive higher reconciliation payments. The stop-loss limit increased from 0% (no repayment responsibility) to 10% over the same period, providing greater repayment risk. The quality-adjusted target price decreased for 94% of mandatory CJR hospitals in PY3 as it became more heavily weighted to the regional average, requiring hospitals to make greater reductions in episode payments to receive a reconciliation payment (Section II.C.1). Finally, as noted above, Medicare began covering TKAs provided in a hospital outpatient setting at the start of PY3.

The CJR model resulted in savings in PY2, and likely resulted in savings in PY1 and PY3 (Exhibit 21). For mandatory CJR hospitals in PY1, Medicare savings is due to a reduction in per-LEJR (IP+OP) non-standardized paid amounts of $1,365, less per-LEJR (IP+OP) reconciliation payments of $819. This resulted in estimated Medicare savings of $546 per LEJR (IP+OP), ranging from losses of up to $17 to savings up to $1,110.

Exhibit 21: The CJR model resulted in Medicare savings for mandatory CJR hospitals in PY2, and likely resulted in savings in PY1 and PY3*

Savings estimates based on all LEJRs (IP+OP), which underestimate the impact of the CJR model

● Reduction in Paid Amounts   ● Medicare Savings

Source: CJR evaluation team analysis of Medicare claims and enrollment data for episodes initiated in 2012 through 2014 that ended between April 2012 and March 2015 (baseline) and LEJRs initiated during or after April 2016 that ended by December 2018 (intervention) and CJR payment contractor data for CJR participant hospitals in PY1-3.
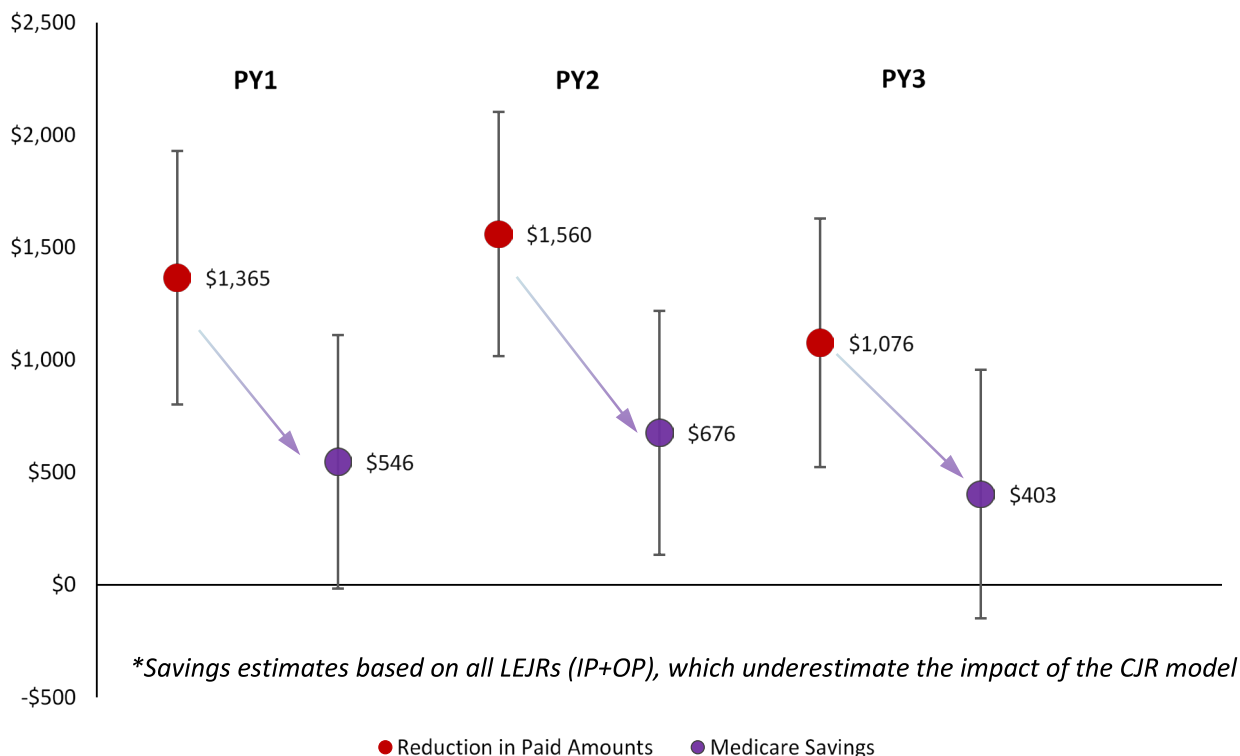
Notes: Ranges based on 90% CIs are plotted as gray bars for reductions in non-standardized paid amounts and Medicare savings. Change in non-standardized paid amounts and CIs are based on estimates from a DiD model of per-LEJR standardized paid amounts that have been multiplied by negative one and converted to non-standardized amounts. Because all LEJRs (IP+OP) are included, Medicare savings are underestimated in PY3.

The arrows represent subtracting average reconciliation payments per LEJR from the reduction in paid amounts to calculate Medicare savings. Negative savings reflect Medicare losses. Average reconciliation payments were computed by dividing total reconciliation payments for CJR episodes by the total number of LEJRs (IP+OP) at mandatory CJR hospitals.

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

For mandatory CJR hospitals in PY2, Medicare savings is due to a reduction in per-LEJR non-standardized paid amounts of $1,560, less per-LEJR (IP+OP) reconciliation payments of $884 (Exhibit 21). This resulted in estimated Medicare savings of $676 per LEJR (IP+OP), ranging from savings of $133 to $1,218.

For mandatory CJR hospitals in PY3, Medicare savings is due to a reduction in per-LEJR non-standardized paid amounts of $1,076, less per-LEJR (IP+OP) reconciliation payments of $673 (Exhibit 21). This resulted in estimated Medicare savings of $403 per LEJR (IP+OP), ranging from losses of up to $150 to savings up to $956.

Savings estimates based on inpatient LEJR episodes

For inpatient LEJR episodes at mandatory CJR hospitals, the CJR model reduced Medicare program spending by an estimated $76.3 million during the first three performance years (Exhibit 22). After accounting for the uncertainty in the per-episode reduction in payments, the estimated savings due to the CJR model ranged from savings of $16.0 million to $136.7 million. Medicare savings is based on an estimated reduction in per-episode non-standardized paid amounts of $1,498 for the 111,351 LEJR that were performed at CJR participant hospitals on an inpatient basis or $166.8 million. From this, we subtracted reconciliation payments of $813 per episode for all 111,351 episodes or $90.5 million, resulting in an estimated savings of $685 per episode, ranging from savings of $143 to $1,227, or in aggregate, savings of $76.3 million ranging from $16.0 million to $136.7 million. The point estimate equates to savings of 2.5% of the baseline.

Exhibit 22: The CJR model resulted in savings to Medicare for mandatory CJR hospitals in PY1-3 (overestimated)

Component	Per episode average		Number of episodes	Total	
	Estimate	Range		Estimate	Range
Reduction in non-standardized paid amounts for inpatient LEJR episodes	$1,498	$956 to $2,040	111,351	$166,785,687	$106,428,007 to $227,143,367
Reconciliation payments	$813	--	111,351	$90,476,027	--
Medicare savings	$685	$143 to $1,227	111,351	$76,309,660	$15,951,980 to $136,667,340

Source: CJR evaluation team analysis of Medicare claims and enrollment data for episodes initiated in 2012 through 2014 that ended between April 2012 and March 2015 (baseline) and episodes initiated during or after April 2016 that ended by December 2018 (intervention) and CJR payment contractor data for CJR participant hospitals in PY1-3.

Notes: Reductions in non-standardized paid amounts are based on a DiD model of per-episode standardized paid amounts that have been multiplied by negative one and converted to non-standardized amounts. Ranges are based on 90% confidence intervals. Estimates are based on inpatient LEJR episodes only and consequently overestimate the true impact of the CJR model.

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

We also evaluated Medicare savings based on inpatient LEJR episodes separately for each performance year to better understand the effect of model features that changed across performance years.

Based on inpatient LEJR episodes, the CJR model resulted in savings in PY2 and PY3 and likely resulted in savings in PY1 (Exhibit 23). For mandatory CJR hospitals in PY1, Medicare savings is due to a reduction in per-episode non-standardized paid amounts of $1,365, less per-episode reconciliation payments of $819. This resulted in estimated Medicare savings of $547 per episode, ranging from losses of up to $16 to savings up to $1,110.

Exhibit 23:   The CJR model resulted in Medicare savings for mandatory CJR hospitals in PY2 and PY3, and likely resulted in savings in PY1*

*Savings estimates based on inpatient LEJRs, which overestimate the impact of the CJR model

● Reduction in Paid Amounts   ● Medicare Savings

Source: CJR evaluation team analysis of Medicare claims and enrollment data for episodes initiated in 2012 through 2014 that ended between April 2012 and March 2015 (baseline) and episodes initiated during or after April 2016 that ended by December 2018 (intervention) and CJR payment contractor data for CJR participant hospitals in PY1-3.

Notes: Ranges based on 90% CIs are plotted as gray bars for reductions in non-standardized paid amounts and Medicare savings. Change in non-standardized paid amounts and CIs are based on estimates from a DiD model of per-episode standardized paid amounts that have been multiplied by negative one and converted to non-standardized amounts. Because only inpatient LEJR episodes are included, Medicare savings are overestimated in PY3.

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

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

For mandatory CJR hospitals in PY2, Medicare savings is due to a reduction in per-episode non-standardized paid amounts of $1,558, less per-episode reconciliation payments of $884 (Exhibit 23). This resulted in estimated Medicare savings of $674 per episode, ranging from savings of $130 to $1,219.

For mandatory CJR hospitals in PY3, Medicare savings is due to a change in per-episode non-standardized paid amounts of $1,490, less per-episode reconciliation payments of $730 (Exhibits 23). This resulted in estimated Medicare savings of $761 per episode, ranging from savings of $146 to $1,376.

d. Conclusion

During the first three performance years, Medicare realized savings for mandatory hospitals in the CJR model. Although the savings estimate varied depending on whether the analysis considered the differential outpatient TKA rate, the cumulative savings estimates were similar. Considering the differential outpatient TKA rate, estimated Medicare savings from the CJR model for all LEJRs (IP+OP) was $61.6 million, ranging from $2.3 million to $120.9 million. While the estimate based on all LEJRs (IP+OP) likely underestimates Medicare savings, it is more representative of the true impact of the CJR model because it considers the differential outpatient TKA rate. Estimated savings for *inpatient LEJR*s which does not consider the impact of the TKA payment policy and overestimates savings, was $76.3 million, ranging from savings of $16.0 million to $136.7 million.

Differences between the savings estimates for *all LEJRs (IP+OP)* and *inpatient LEJRs* were evident for PY3, suggesting that the removal of TKA from the inpatient only list reduced Medicare savings attributable to the CJR model. As expected, the savings estimates for *all LEJRs (IP+OP)* and *inpatient LEJRs* are nearly identical for PY1 and PY2. They indicate that the CJR model likely resulted in Medicare savings in PY1 and did result in Medicare savings in PY2.

The Medicare savings estimate based on all LEJRs and the estimate based on inpatient LEJRs only bound the true savings impact of the CJR model at mandatory CJR hospitals, although the estimate for all LEJRs is closer to the truth. This is because, as indicated by the reduced savings in PY3, the estimate from all LEJRs (IP+OP) considers the financial incentives under the CJR model to keep LEJRs in the more expensive inpatient setting. These estimates do not, however, account for the full impact of the CJR model because they do not include episode payment changes achieved by all of the hospitals that ever participated in the CJR model or their reconciliation payments. In future reports, we will estimate Medicare savings for the full expression of the CJR model and evaluate how changes to model features and Medicare policies affected Medicare savings due to the CJR model.

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

Changes in service-level payments and use provide insights into how hospitals reduced average episode payments. Hospital and PAC payments comprise the largest shares of average episode payments with each accounting for about 40%. Prior to the OP TKA coverage change, hospital payments were unlikely to change because hospitals receive a per-discharge payment for inpatient LEJRs that typically is not affected by length of stay or services provided during the hospitalization. The Medicare payment for a TKA performed in the outpatient setting, however, is lower than the inpatient payment. So, relative hospital anchor payments may differ if there is differential distribution of inpatient versus outpatient procedures. PAC payments can be reduced by shifting service use from more to less expensive care settings. Generally, average IRF payments are higher than average SNF payments, and Medicare payments for both of these institutional PAC settings tend to be higher than payments for home health (HH) care. An additional effect of OP

TKA coverage could be lower PAC payments because Medicare does not cover SNF stays following an outpatient TKA.

This section examines changes in care pathways following an inpatient LEJR and therefore presents impact estimates based on *inpatient LEJR episodes*. These estimates likely overestimate the impact of the CJR model because, as described briefly above and in more detail in Section II.A.2, they do not reflect the impact that the CJR model had on the decision about whether to provide TKA in the inpatient or outpatient setting. The impact estimates based on *all LEJRs (IP+OP)* are provided in Appendix B (Exhibit B-3). While magnitudes differ, both sets of estimates indicate that the reduction in average episode payments was due to reductions in institutional PAC.

---

a. Key Findings

■ The relative decrease in average episode payments was driven by relative decreases in institutional PAC payments.

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

b. Methods

This analysis uses a DiD design (described in Section II.B.1) 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.[34] Average payments by service are calculated across all episodes, including episodes in which the patient did not receive the particular service.

c. Results

During the first three performance years, the relative decrease in average inpatient episode payments ($1,540, p<0.01) was driven by relative reductions in IRF and SNF payments. There were also relative reductions in readmission and Part B payments. Average IRF payments decreased by $539 more for CJR inpatient episodes than for control group episodes from the baseline to the intervention period (p<0.01, Exhibit 24). This equates to a 24.4% decrease in average IRF payments for CJR episodes from the baseline. Average SNF payments decreased by $935 more for CJR inpatient episodes than for control group episodes, or 15.4% from the CJR baseline (p<0.01, Exhibit 24). Average readmission payments decreased by $140 more for CJR inpatient episodes than for control group episodes, or 11.6% from the CJR baseline (p<0.1, Exhibit 24). Average Part B payments decreased by $104 more for CJR inpatient episodes than for control group episodes, or 2.1% from the CJR baseline (p<0.1, Exhibit 24).[35]

---

[34]   Appendix F – Outcome Definitions
[35]   Part B payments include payments for physician and non-physician practitioners, hospital outpatient departments, ambulatory surgical centers, and durable medical equipment.

Exhibit 24:   The reduction in average episode payments was driven by decreases in inpatient rehabilitation facility and skilled nursing facility payments, PY1-3*

*Impact estimates based on inpatient LEJRs, which overestimate the impact of the CJR model

Source: CJR evaluation team analysis of Medicare claims and enrollment data for episodes initiated in 2012 through 2014 that ended between April 2012 and March 2015 (baseline) and episodes initiated during or after April 2016 that ended by December 2018 (intervention).

Notes: The estimates in this exhibit are the result of a DiD model. DiD estimates that are significant at the 99%, 95%, or 90% significance level are indicated by red, orange, or yellow shaded bars, respectively. An inpatient LEJR episode begins with an inpatient anchor hospitalization that meets CJR episode eligibility requirements and ends 90 days after discharge. Inpatient LEJR episodes are CJR episodes under the CJR model.

Impact estimates based on inpatient LEJR episodes overestimate the impact of the CJR model.

⚠ Results need to be interpreted with caution because data from the baseline period showed CJR and control group hospitals were not on parallel trends for these outcomes, which is required for an unbiased estimate.

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

Exhibit 25: The reasons for the decrease in post-acute care payments differ by setting*
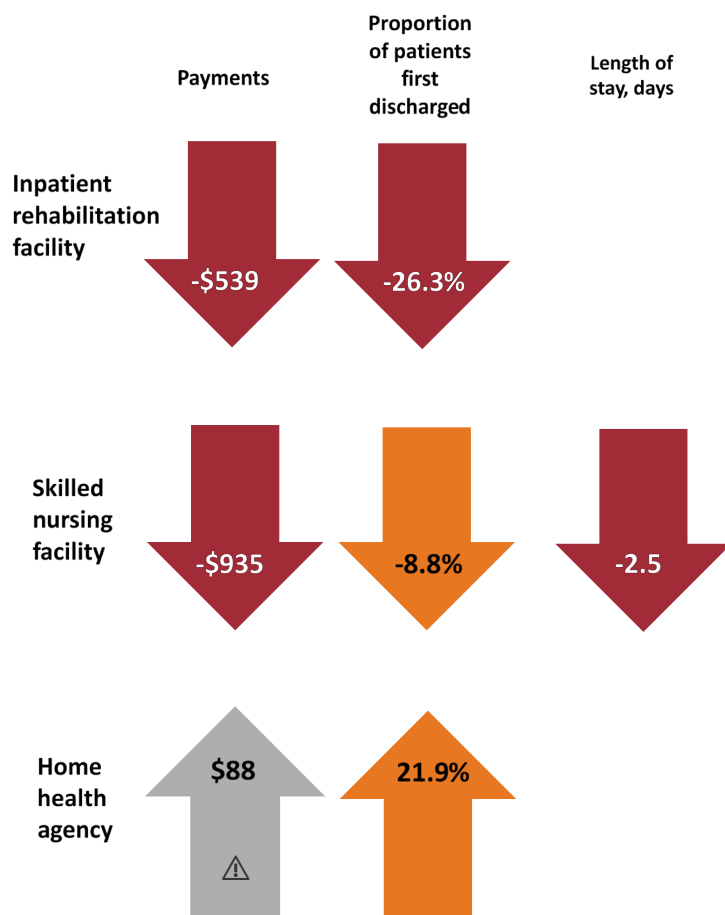
Impact estimates based on inpatient LEJRs, which overestimate the impact of the CJR model

Source: CJR evaluation team analysis of Medicare claims and enrollment data for episodes initiated in 2012 through 2014 that ended between April 2012 and March 2015 (baseline) and episodes initiated during or after April 2016 that ended by December 2018 (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 red, orange, or yellow shaded shapes, respectively.

An inpatient LEJR episode begins with an inpatient anchor hospitalization that meets CJR episode eligibility requirements and ends 90 days after discharge. Inpatient LEJR episodes are CJR episodes under the CJR model.

Impact estimates based on inpatient LEJR episodes overestimate the impact of the CJR model. The change in the proportion of patients first discharged to each PAC setting represents the percent change from the CJR baseline that is due to CJR. It is calculated by dividing the DiD estimate by the CJR baseline average.

For SNF LOS, beneficiaries must have spent at least one day in the SNF setting. ⚠ Results need to be interpreted with caution because data from the baseline period showed CJR and control group hospitals were not on parallel trends for these outcomes, which is required for an unbiased estimate. DiD = difference-in-differences, LEJR = lower extremity joint replacement, LOS = length of stay, PAC = post-acute care, SNF = skilled nursing facility.

The relative decrease in IRF payments is the result of a relative reduction in the proportion of CJR patients discharged from the hospital to an IRF. The proportion of patients first discharged to an IRF decreased more for CJR inpatient episodes than for control episodes, representing a 26.3% decrease from the CJR baseline proportion (p<0.01, Exhibit 25).

During the baseline, 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 the decline was greater for CJR patients (Exhibit 26). There was no change in the number of days that CJR patients spent in an IRF, among patients with at least one IRF stay (Appendix B). This is not unexpected because reducing LOS does not affect Medicare's per discharge payment.

The relative decrease in SNF payments is due to a relative reduction in the proportion of CJR patients discharged from the hospital to a SNF and a relative decrease in the number of SNF days among patients with at least one SNF stay. During the

baseline, a smaller proportion of CJR patients were discharged to a SNF than control group patients. The proportion of patients discharged to a SNF decreased more for CJR inpatient episodes than for control group episodes from the baseline to the intervention period, representing an 8.8% decrease from the CJR baseline (p<0.05, Exhibit 25 and 27). For patients with a SNF stay, the average number of SNF days decreased by 2.5 days more for CJR than for control group episodes from the baseline to the intervention period (p<0.01, Exhibit 25). Medicare pays SNFs a daily rate for each day of care, so fewer SNF days result in lower episode payments.

We cannot draw a conclusion about whether the CJR model affected HHA payments. Although our estimate indicates there was no relative change in average HHA payments, multiple statistical tests indicate that CJR and control group hospitals did not have parallel trends for this outcome during the baseline period, so the assumptions of the statistical model are violated. This means that we are not able to conclude that there is indeed no change in HHA payments as a result of the CJR model. There were changes in HHA use. The proportion of patients first discharged to an HHA increased more for CJR inpatient episodes than for control group episodes. The proportion of CJR patients first discharged to an HHA increased 21.9% from the CJR baseline proportion (p<0.05, Exhibits 25 and 28).

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

Exhibit 26:   The proportion of patients discharged to inpatient rehabilitation facilities declined more for CJR than for control episodes, PY1-3*

***Source**:   CJR evaluation team analysis of Medicare claims and enrollment data for episodes initiated on or after January 2012 that ended by December 2018.*

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

An inpatient LEJR episode begins with an inpatient anchor hospitalization that meets CJR episode eligibility requirements and ends 90 days after discharge. Inpatient LEJR episodes are CJR episodes under the CJR model.

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

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

Exhibit 27:   The proportion of patients discharged to skilled nursing facilities declined more for CJR than for control episodes, PY1-3*

Source: CJR evaluation team analysis of Medicare claims and enrollment data for episodes initiated on or after January 2012 that ended by December 2018.

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

An inpatient LEJR episode begins with an inpatient anchor hospitalization that meets CJR episode eligibility requirements and ends 90 days after discharge. Inpatient LEJR episodes are CJR episodes under the CJR model.

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

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

Exhibit 28: The proportion of patients first discharged to home health agencies increased more for CJR episodes than for control episodes, PY1-3*

Source: CJR evaluation team analysis of Medicare claims and enrollment data for episodes initiated on or after January 2012 that ended by December 2018.

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

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

An inpatient LEJR episode begins with an inpatient anchor hospitalization that meets CJR episode eligibility requirements and ends 90 days after discharge. Inpatient LEJR episodes are CJR episodes under the CJR model.

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

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

The relative decrease in average readmission payments was primarily due to changes made by a few CJR participant hospitals, rather than a reduction in the overall readmission rate. In the baseline, 59 CJR participant hospitals discharged 1,640 LEJR patients to IPPS hospitals for rehabilitation (under MS-DRG 945 or 946), which represents 1.2% of baseline episodes. Ninety percent of these patients were discharged to five hospitals. In the control group, 19 hospitals discharged 0.02% of LEJR patients to IPPS hospitals for rehabilitation during the baseline. In the intervention period this relatively uncommon practice declined further so that 32 CJR and 9 control group patients were discharged to an IPPS hospital for rehabilitation, resulting in the decrease in readmission payments.

d.   Conclusion

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

SNF and HHA Medicare Payment Changes

On October 1, 2019, CMS changed the SNF and HHA payment systems, which will likely affect SNF and HHA utilization and payments. The SNF Patient-Driven Payment Model (PDPM) and the Home Health Patient-Driven Groupings Model (PDGM) both use clinical and functional characteristics of patients rather than volume of therapy services furnished to classify patients for payment, which reduces incentives to provide more therapy to boost payments. In addition, as of October 1, 2018, Medicare began adjusting SNF payments based on their 30-day All-Cause Readmission Measure. Most PAC provider interviewees who discussed the payment changes did not anticipate a significant financial impact; a few indicated that the changes result in more appropriate accounting for patient characteristics and needs. In the future, we will investigate whether and how these payment reforms affect CJR participant hospitals differently than non-CJR hospitals.

---

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

The CJR model was designed to reward hospitals that delivered high quality of care for Medicare beneficiaries undergoing LEJRs. 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.

a. Key Findings

- The unplanned readmission rate improved under the CJR model, while emergency department use and the mortality rate were maintained.

- For inpatient elective LEJR episodes, there was a reduction in the complication rate. For TKA episodes, there was no statistically significant change in the manipulation under anesthesia (MUA) rate.

b. Methods

This analysis used the DiD approach described in Section II.B.1 to estimate the relative change in outcomes. Appendix E (Section III.D.1) includes more detailed information about the methodology.

c. Results

During the first three performance years, the CJR model resulted in a relative reduction in the 90-day unplanned readmission rate and had no statistically significant impact on the 90-day emergency department (ED) use rate or the mortality rate during the anchor hospitalization plus the 90-day post-discharge period for inpatient LEJR episodes (Exhibit 29). Results for all LEJR episodes (including OP TKA episodes) were generally similar and are in Appendix B (Exhibit B-3).

The unplanned readmission rate decreased more for CJR episodes than for control group episodes, representing a 3.1% decrease from the CJR baseline ($p < 0.10$).

While the change in mortality was not statistically significantly different between mandatory CJR and control hospitals, it did represent a 2.0% relative increase for CJR episodes from the CJR baseline. Because of the importance of the measure, we further examined the estimated relative increase. For mandatory CJR hospitals, the mortality rate decreased from 2.6% in the baseline to 2.5% in the intervention period. For control hospitals, it decreased from 2.8% to 2.6% during the same time (Appendix B, Exhibit B-1). The control group had a higher baseline mortality rate and during the intervention period decreased the mortality rate to approximately the same level as that of the CJR group. The mortality rates are comparable during the intervention period.

Exhibit 29:   Quality of care was maintained or improved under the CJR model for mandatory CJR hospitals, PY1-3*

-3.1%   Unplanned readmission rate

0.2%   ED use

2.0%   Mortality rate

-7.4%   Complication rate, elective LEJR

1.1%   Manipulation under anesthesia, TKA

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

Relative change in quality metric from baseline

*Impact estimates based on inpatient LEJRs, which overestimate the impact of the CJR model

Source: CJR evaluation team analysis of Medicare claims and enrollment data for episodes initiated in 2012 through 2014 that ended between April 2012 and March 2015 (baseline) and episodes initiated during or after April 2016 that ended by December 2018 (intervention).

Notes: The estimates in this exhibit are the result of a DiD model. DiD estimates that are significant at the 99%, 95%, or 90% significance level are indicated by red, orange, or yellow shaded bars, respectively.

An inpatient LEJR episode begins with an inpatient anchor hospitalization that meets CJR episode eligibility requirements and ends 90 days after discharge. Inpatient LEJR episodes are CJR episodes under the CJR model. Impact estimates based on inpatient LEJR episodes overestimate the impact of the CJR model.

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

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

> Manipulation under anesthesia (MUA) is performed to treat stiffness and improve range of motion after TKA. Members of our clinical review panel suggested that because institutional PAC use decreased under the CJR model (Section II.B.3), it would be valuable to evaluate the MUA rate. They indicated that an increase in the MUA rate might signal inadequate rehabilitation under the CJR model.

We also evaluated changes in the complication rate, which is specific to elective LEJR, and the MUA rate, which is specific to elective TKA. The complication rate for elective LEJR decreased more for CJR episodes than control group episodes, representing a 7.4% decrease from the CJR baseline.[36]

The CJR model had no statistically significant impact on the MUA rate in the 90-day post-discharge period for TKA. While it was not statistically significant, we also further investigated the estimated 1.1% relative increase in the MUA rate. For mandatory CJR hospitals, the MUA rate stayed constant at 1.9% from the baseline to the intervention period. For control group hospitals, the rate decreased from 2.0% to 1.9% during the same time (Appendix B, Exhibit B-1). Thus, the MUA rate was equivalent for mandatory CJR and control group hospitals during the intervention period.

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

The CJR model is intended to encourage participant hospitals to improve quality by coordinating care with all providers and clinicians involved in the episode, however, the incentive to lower episode payments could lead to changes in care that result in worse long-term outcomes. For this reason, functional status and pain in the period after the procedure and at later stages of recovery are of particular importance for understanding the impact of the CJR model on quality of care.

Functional status, pain, and care experience are indicators of quality of care that cannot be assessed with Medicare claims data. We therefore surveyed patients after the end of their episodes to learn about their experiences and recovery. We also examined patient assessments conducted while patients were receiving PAC to understand interim measures of function and pain.

The type and duration of care that a patient receives after LEJR could affect functional recovery and pain. Under the CJR model, fewer patients were discharged to institutional PAC settings and more were discharged home with HH care. For CJR patients who used a SNF, the number of SNF days decreased from the baseline to the intervention period, relative to control patients (Section II.B.3). It is therefore important to consider not only the overall impact of the CJR model on functional status and pain, but also for patients receiving care in the different PAC settings.

---

[36] For all elective LEJRs (IP+OP), the complication rate decreased more for mandatory CJR hospitals than for control group hospitals, but the relative decrease was not statistically significant (-6.5% from the CJR baseline, p=0.10, Appendix B, Exhibit B-3).

> ## a. Key Findings
>
> ■ After the end of the episode, CJR and control survey respondents had similar self-reported changes in functional status and pain.
>
> ■ CJR and control patients first discharged to an IRF or an HHA had similar improvement in functional status and pain while in the PAC setting. For patients first discharged to a SNF, one measure indicated that relatively fewer CJR patients improved in functional status, and the measure of pain indicated a relative improvement for CJR patients.
>
> ■ CJR and control survey respondents reported similar satisfaction with overall recovery and care management.
>
> ■ Overall slightly fewer CJR survey respondents reported satisfactory care transition experiences, and a relatively small number of CJR respondents required more caregiver help at home than control respondents. Results were driven by CJR respondents first discharged to an IRF or a SNF. CJR respondents first discharged to an IRF were more complex than control respondents.

b. Methods

Patient survey

We surveyed beneficiaries after the end of their inpatient LEJR episode to determine if CJR patients differed from control patients on several patient-reported outcomes.[37] Measures included change in functional status and pain (recalled from before their surgery to the time of the survey), as well as satisfaction with overall recovery, satisfaction with care management, experience with care transitions, and caregiver help needed after returning home. We estimated risk-adjusted differences between CJR and control respondents, accounting for beneficiary, hospital, and MSA attributes. See Appendix E (Section VI) for more detail on the methods and Appendix G for the patient survey questions.

Changes in functional status and pain were measured with a Likert scale that captures distinct levels of function and pain. We transformed differences in levels into percentage terms, where the numerator is the estimated difference in change (from before their surgery to after the end of the episode), and the denominator is the average recalled functional status among CJR respondents from before their surgery. We report differences in the change in functional status in levels and percentage terms. While this provides an indication of the average change, it does not distinguish between changes that are large and concentrated among fewer respondents and changes that are small and spread across more respondents. For example, an estimated difference of -0.04 for a measure with five response categories could indicate that 4 out of every 100 respondents differed by one category of function, or that 1 out of every 100 respondents differed by four categories of

---

[37] The median time at which surveys were returned was 35 days after the conclusion of the patient's 90-day post-discharge period.

function, or something in between. Measures of satisfaction and caregiver help were scaled from 0-100 points. Measures of experience with care transitions are reported in percentage terms (with differences interpreted on a percentage point scale).

Data were collected in two waves that covered episodes with inpatient discharges from mandatory CJR hospitals in March or April 2018 and in August or September 2018. We included all CJR episodes initiated during these periods and randomly selected control episodes, which represented 89.3% of all control episodes, to obtain a similar sample size to the CJR sample.[38]

The overall response rate was 67.6% in the CJR group and 68.6% in the control group. Survey results were based on 8,433 completed survey responses from CJR patients and 9,014 from control patients. In aggregate, CJR and control respondents had similar patient characteristics (Appendix E Section VI.F).

In addition to analyzing episodes for all patients, we analyzed subgroups based on the first PAC setting to which patients were discharged from the hospital. This analysis explores whether functional status and care experiences varied by PAC setting, since those who used an IRF or a SNF may have had different experiences than those who did not require institutional PAC. Exhibit 30 below shows sample sizes and response rates by setting. CJR patients who were first discharged to an IRF or a SNF were significantly less likely to respond to the survey than control respondents. This is consistent with greater patient complexity among CJR patients in these settings, as more complex patients are less likely to respond to surveys. Summary statistics of key variables by setting are detailed in Appendix E, Section VI.F.

Exhibit 30: Patient survey sample size and response rate overall and by discharge setting

Group	Patients surveyed (starting sample)[a]		Survey responses received (analytic sample)		Response rate	
	CJR	Control	CJR	Control	CJR	Control
Overall	12,478	13,137	8,433	9,014	67.6%	68.6%
Home without HH	2,001	2,511	1,485	1,852	74.2%	73.8%
HHA	6,193	5,760	4,584	4,256	74.0%	73.9%
SNF	3,440	3,760	1,881	2,243	54.7%	59.7%**
IRF	736	1,042	398	620	54.1%	59.5%*

Source: CJR evaluation team analysis of survey data for patients with discharge from LEJR surgery in March, April, August, or September 2018.

Notes: Significance of difference in response rate determined by t-test: *p<0.10 ** p < 0.05.
HH = home health, HHA = home health agency, IRF = inpatient rehabilitation facility, LEJR = lower extremity joint replacement, SNF = skilled nursing facility.

[a] The starting sample and analytic sample do not sum to the same values as Exhibit E-19 (Appendix E) because data on post-acute care discharge were missing for about 1% of observations.

---

[38] We initially drew an equal number of control episodes. After the sample was pulled, CMS identified 7 CJR participant hospitals that were retroactively identified as "rural" and excluded from mandatory participation; these CJR patients were dropped from the sample.

We cannot be certain that any differences in patient survey results by discharge setting were caused by the CJR model. The MSA-level randomization in the CJR model ensured a balanced control group overall, but not necessarily within all PAC settings (Appendix E, Exhibits E-24 and E-26). After the CJR model began, fewer CJR patients were first discharged to an IRF or a SNF and more were discharged to an HHA, relative to the control group (Section II.B.3). With these shifts, we would expect patient complexity to increase in each PAC setting for CJR patients.[39] Although the stratified analysis adjusts for observable patient characteristics within each setting, decisions regarding PAC setting may be based, in part, on patient characteristics that cannot be observed in the data (such as presence of a caregiver in the home, or health conditions not included in the HCC score). To the extent that CJR patients require more resources for reasons we cannot observe, estimates will be biased downward (appearing to indicate worse outcomes for CJR patients). Stratifying results by PAC setting establishes a lower-bound on the estimate of any negative impact of the CJR model within each setting.[40]

Assessment-based measures

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

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

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

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

The measures differ across these setting-specific instruments, as does the timing between the admission and discharge assessments.[41] While this precludes direct comparisons of patients across settings, the measures provide valuable information about changes in functional status and pain while a patient is receiving care in the setting. We focus on the first PAC setting to which patients were discharged (IRF, SNF, or HHA) and measure changes in patient functional status and pain. We report one functional status measure for patients initially discharged to an IRF, two functional

---

[39] For example, if complex patients who may have been discharged to an IRF prior to the CJR model were discharged to a SNF under the CJR model, patient complexity would increase in both settings on average.

[40] This is true even for positive estimated differences. For example, if we estimated that the CJR model improved a measure of function by 5% relative to the control group, this improvement would still be the lower bound of the impact of the CJR model, and thus a "worst case scenario". In such a case, the true impact of the CJR model may be greater than 5%.

[41] The IRF-PAI, MDS, and OASIS assessments use different instruments to assess functional status throughout the stay. Even though in October 2018, CMS added cross-setting measures of functional status to each of the three assessment instruments, we are not using the new measures in this analysis because this report does not cover assessment data after September 30, 2018.

status measures and one pain measure for those initially discharged to a SNF, and two functional status measures and one pain measure for those initially discharged to an HHA.

The results are risk-adjusted to control for functional status at the initiation of PAC with health care service use before the anchor hospitalization and beneficiary, market, and hospital characteristics.[42] Risk adjusting is important because the CJR model has affected the initial discharge setting for LEJR patients, as well as the duration of PAC in each setting.

We conducted additional analyses to assess how the complexity of CJR patients discharged to an IRF, SNF, or HHA changed relative to 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. Details about the DiD estimator and risk adjustment models are discussed in Appendix E (Section III.D.1 and Section III.D.2).

The assessment-based analyses rely on the same baseline period as the claims-based analyses. The intervention period is one quarter shorter (from April 2016 to September 2018) because of the longer time needed for PAC assessment data to become available.[43] Results from inpatient episodes are reported in this chapter; results for all LEJRs (IP+OP) were generally similar and reported in Appendix B (Exhibit B-4).

Interpreting survey and assessment results by first discharge setting

Results from the assessments and the patient survey may differ because the assessment measures generally cover a period within a few weeks of discharge, while the survey captures patients' status after the episode has ended, and patients' short-term and long-term outcomes may differ. There could also be differences in the two results due to differences in perspective, since patient survey data are self-reported, while assessment data are reported by a clinician. Differences could also be due to differences in the measure definitions.

c. Results

Overall

Patient survey results

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

Functional status and pain

Overall, CJR and control respondents reported similar improvements from before their surgery to after the episode on all eight measures of functional status and pain (Exhibit 31). Differences in the

---

[42] 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.

[43] 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.

amount of improvement between CJR and control respondents, as a percentage of the average status CJR respondents recalled prior to their hospitalization, were generally not significantly different. The only statistically significant result was CJR respondents reported less improvement in rising from sitting than control respondents (-1.1%, p<0.10).

Exhibit 31:   CJR and control survey respondents experienced similar improvement in functional status and pain

Survey measure	Response range	Mean change in self-reported measure from before the hospitalization to after the episode *Higher value represents a more favorable change*		Difference between CJR and control groups (% difference)[a]
		CJR	Control group	
Ability to walk by yourself without resting	-4 to 4	0.77	0.75	0.02 (0.7%)
Difficulty walking up or down 12 stairs	-3 to 3	0.75	0.75	-0.01 (-0.4%)
Difficulty rising from sitting	-4 to 4	1.21	1.24	-0.03 (-1.1%)
Difficulty standing	-4 to 4	1.16	1.17	-0.01 (-0.3%)
Use of a mobility aid	-2 to 2	0.15	0.16	-0.01 (-0.4%)
Difficulty getting on/off the toilet	-4 to 4	1.37	1.39	-0.01 (-0.5%)
Frequency that pain interferes with normal activities	-4 to 4	1.96	1.98	-0.02 (-1.0%)
Medication use for pain in the joint you had replaced	-3 to 3	0.61	0.61	0.00 (-0.1%)

Source: CJR evaluation team analysis of patient survey data for episodes with discharge in March, April, August, or September 2018.

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

Survey sample sizes are reported in Exhibit 30.

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

Satisfaction, experience with care transitions, and caregiver help[44]

The patient survey also asked about satisfaction with overall recovery, satisfaction with care management, experience with care transitions, and caregiver help needed after returning home.

Satisfaction

There were no statistically significant differences between CJR and control respondents on any measure of satisfaction related to overall recovery or care management (Appendix H).

---

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

Experience with care transitions

Both the CJR and control respondents generally indicated positive experiences with care transitions, with over 80% in both groups reporting they were discharged at the right time, received the right amount of PAC, and had access to all necessary durable medical equipment (Exhibit 32). However, roughly one fewer CJR respondent out of every 100 reported that they were discharged at the right time (1.1 percentage points, p<0.10) or received the right amount of PAC in the two weeks following discharge (1.2 percentage points, p<0.05).

Exhibit 32:   Care transition experiences were generally positive, but slightly fewer CJR respondents reported satisfactory care transition experiences than control respondents
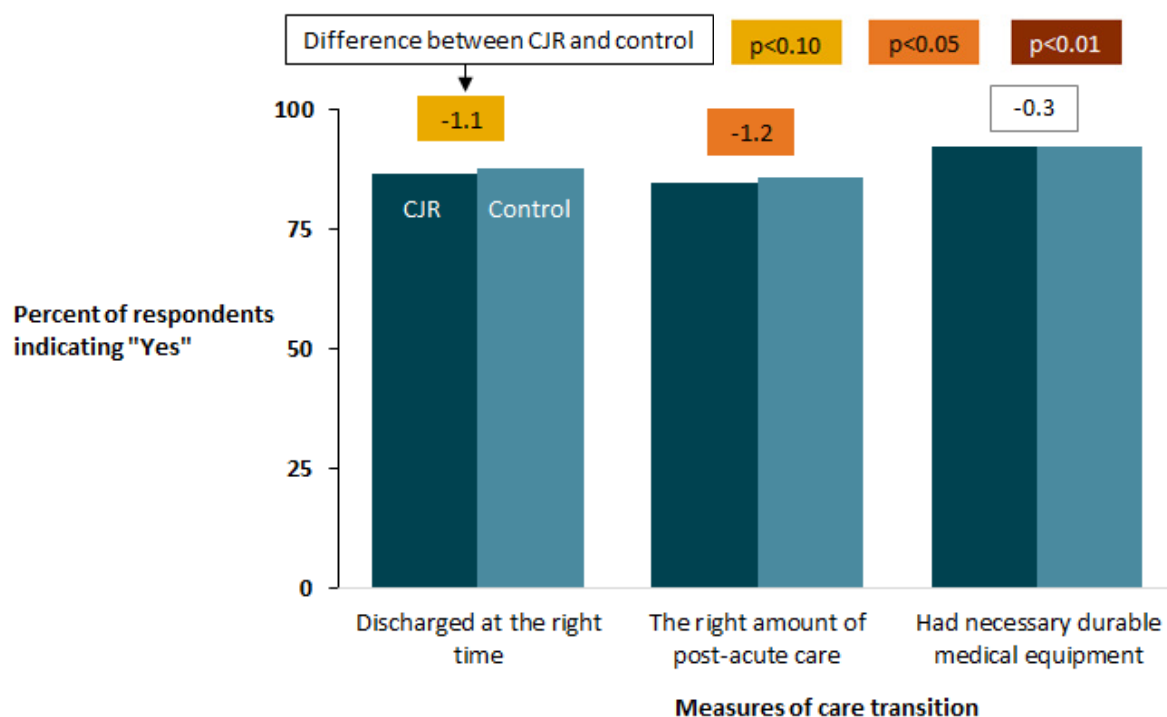
Source: CJR evaluation team analysis of patient survey data for episodes with discharge in March, April, August, or September 2018.

Notes: Differences that are significant at the 99%, 95%, or 90% significance level are indicated by red, orange, or yellow shading, respectively.

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

Survey sample sizes are reported in Exhibit 30.

Caregiver help

Approximately 95% of CJR and control survey respondents received help from a caregiver after returning home, and this proportion was not statistically significantly different between the two groups (Appendix H). Among those who received help from a caregiver, CJR respondents needed more help than control respondents, based on a composite measure of help with three activities of daily living (ADLs) (-1.9 points out of 100, p<0.01, Exhibit 33). All three ADL measures contributed to the difference in the composite score, with CJR respondents needing more help than

control respondents with putting on and taking off clothes (-2.2 points, p<0.01), bathing
(-2.6 points, p<0.01), and using the toilet (-1.7 points, p<0.05).

Exhibit 33:   CJR survey respondents required more help with ADLs from caregivers at home than control respondents
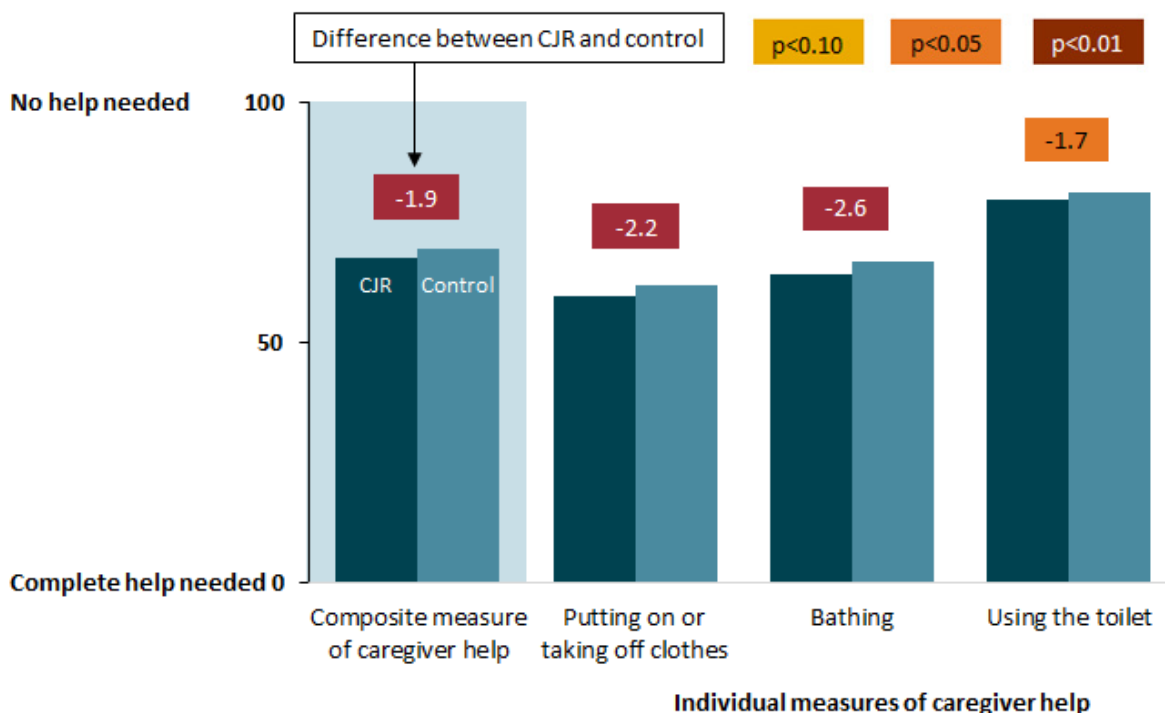
Notes: Differences that are significant at the 99%, 95%, or 90% significance level are indicated by red, orange, or yellow shading, respectively.

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

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

Survey sample sizes are reported in Exhibit 30.

ADLs = activities of daily living.

Results by first PAC setting

The CJR model resulted in shifts in PAC use, which could affect functional recovery and pain. Under the CJR model, relatively fewer patients were first discharged to an IRF or SNF and relatively more patients were discharged to an HHA. While survey results did not indicate aggregate differences between CJR and control respondents in changes in functional status or pain, aggregate results could mask differential impacts based on PAC setting. That is, better changes for respondents discharged to one PAC setting could offset worse changes for those discharged to a different PAC setting. To investigate this possibility, this section presents changes in measures of functional recovery and pain within each PAC setting. As discussed in the methods section (II.B.5.b), results by PAC setting should be interpreted with caution as they

may be driven by an increase in patient complexity in the CJR group relative to the control group. While we cannot be certain that any differences in results by discharge setting were caused by the CJR model, a lack of meaningful differences within each setting would provide evidence that potential CJR impacts in each setting were limited.

Patients first discharged to an IRF

Changes in patient mix

The complexity of the average CJR patient first discharged to an IRF was greater during the intervention period than the baseline period, relative to the control group (Exhibit 34). There were relative increases in the proportion of CJR patients with characteristics that are associated with higher episode costs, suggesting an increase in complexity.[45] There were relative increases in the proportion of CJR patients who were 80 years or older (p<0.01), had dementia (p <0.01), had an acute care hospital stay in the six months preceding their episode (p<0.05), were discharged under the more complex MS-DRG 469 (p<0.05), or had a fracture (p<0.01). Further, there was a relative increase in the average HCC score[46] (p<0.01) and a relative increase in the average mobility index (p<0.05) for CJR patients compared to control group patients.[47] For each of these measures, the average complexity of both CJR and control patients increased from the baseline to intervention, but the increase was greater for CJR patients.

---

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

[46] The HCC score predicts Medicare spending in the coming year for a given Medicare beneficiary compared to the average beneficiary. Beneficiaries with scores above 1.0 are more complex and predicted to have higher costs in the coming year than the average beneficiary and beneficiaries with scores below 1.0 are healthier and predicted to have lower costs than the average.

[47] The IRF-PAI measures the functional status of IRF patients based on the Functional Independence Measure (FIM) scores, which are higher for patients with more functional independence. We multiplied the FIM scores by -1 so they matched the direction of other patient mix measures (i.e., a higher score indicates greater complexity).

Exhibit 34:   Patient complexity increased for CJR patients first discharged to an IRF, PY1-3

Characteristic	Net difference	Lower episode payments	Higher episode payments	p-value
Demographics				
Age, 80+	4.9		●————	<0.01
Sex, Female	-1.8	○		<0.10
Race, Black	-0.9	●		0.32
Social Determinants				
Eligible for Medicaid	0.9	●		0.41
Disability, no ESRD	1.0	●		0.23
Health status				
HCC Score	0.2	●		<0.01
Obesity	-2.5	●		0.11
Diabetes	0.2	●		0.86
Hypertension	-1.1	●		0.20
Dementia	2.4	●		<0.01
Congestive Heart Failure	0.1	●		0.94
Hip Fracture	11.3		●	<0.01
MS-DRG 469	2.1	●		<0.05
Mobility index	0.4	●		<0.05
Self-care index	0.3	●		0.48
Cognitive index	0.1	●		0.81
Prior use				
ACH Stay	1.8	●		<0.05
HH Use	0.4	●		0.76
IRF Stay	0.5	●		0.31
SNF Stay	0.4	●		0.33
Any Prior Care	1.4	●		0.21

-6   -3   0   3   6   9   12   15

Net difference (90% CI)

Source: CJR evaluation team analysis of Medicare claims, enrollment, and 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 December 2018 (intervention).

Notes: Net differences at the 99%, 95%, or 90% significance level are indicated by red, orange, or yellow shaded cells, respectively, and are considered to indicate a relative increase or decrease in patient complexity.

ACH = acute care hospital, CI = confidence interval, ESRD = end-stage renal disease, HCC = hierarchical condition category, HH = home health, IRF = inpatient rehabilitation facility, IRF-PAI = Inpatient Rehabilitation Facility Patient Assessment Instrument, MS-DRG = Medicare Severity-Diagnosis Related Group, PY = performance year, SNF = skilled nursing facility.

Functional status and pain

Patient survey results

Overall, CJR survey respondents first discharged to an IRF experienced less improvement on three of eight measures of functional status and pain than control respondents, from before their surgery to after the end of the episode (Exhibit 35). CJR respondents reported less improvement than control respondents in standing (-5.0%, $p<0.05$), increased dependence on a mobility device

(-3.7%, p<0.10), and less improvement in getting on and off the toilet (-4.0%, p<0.01). These estimates are likely biased downward (appearing to indicate worse outcomes for CJR patients) because patient complexity increased in the IRF setting for CJR patients due to the shift in first PAC setting.

Exhibit 35: **CJR survey respondents first discharged to an IRF reported less improvement in functional status than control respondents**

| Survey measure | Response range | Mean change in self-reported measure from before the hospitalization to after the episode *Higher value represents a more favorable change* | | Difference between CJR and control groups (% difference)[a] |
		CJR	Control group	
Ability to walk by yourself without resting	-4 to 4	0.22	0.31	-0.09 (-3.4%)
Difficulty walking up or down 12 stairs	-3 to 3	0.51	0.50	0.01 (0.4%)
Difficulty rising from sitting	-4 to 4	0.97	1.05	-0.09 (-2.6%)
Difficulty standing	-4 to 4	0.86	1.03	-0.18 (-5.0%)
Use of a mobility aid	-2 to 2	-0.15	-0.07	-0.08 (-3.7%)
Difficulty getting on/off the toilet	-4 to 4	1.11	1.26	-0.14 (-4.0%)
Frequency that pain interferes with normal activities	-4 to 4	1.77	1.80	-0.04 (-1.2%)
Medication use for pain in the joint you had replaced	-3 to 3	0.46	0.50	-0.04 (-1.2%)

Source: CJR evaluation team analysis of patient survey data for patients first discharged to an IRF in March, April, August, or September 2018.

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

Survey sample sizes by setting are reported in Exhibit 30.

IRF = inpatient rehabilitation facility.

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

Assessment-based results

CJR and control respondents had similar improvement in average mobility scores during their IRF stays. Between the baseline and intervention periods, changes in the amount that patients' mobility improved during their IRF stay were not statistically significantly different for CJR and control patients whose initial PAC setting was an IRF (Exhibit 36).[48] Estimates for both the patient assessment measure and the patient-reported survey measures are small in magnitude. This

---

[48] 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.

suggests that differences are attributable to differences between the two sources of data rather than substantive differences between patients' short-term and longer-term functional recovery.

Exhibit 36: **For patients first discharged to an IRF, no statistically significant difference in the change in mobility score for CJR patients and control patients, PY1-3**
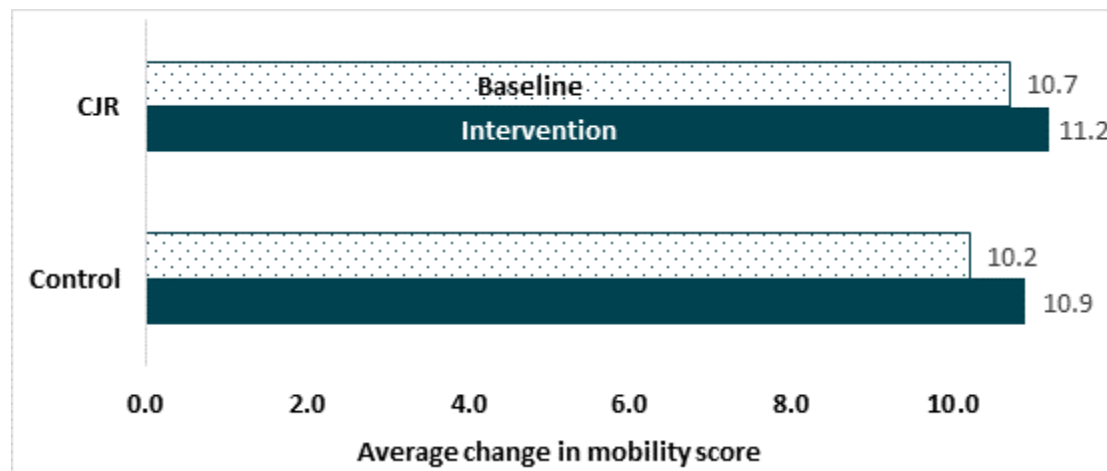
Source: CJR evaluation team analysis of MDS data for episodes initiated in 2012 through 2014 that ended between April 2012 and March 2015 (baseline) and episodes initiated during or after April 2016 that ended by September 2018 (intervention).

Note: IRF = inpatient rehabilitation facility, MDS = Minimum Data Set, PY = performance year.

Satisfaction, experience with care transitions, and caregiver help[49]

Satisfaction

There were no statistically significant differences between CJR and control survey respondents first discharged to an IRF on any measure of satisfaction related to overall recovery or care management.

Experience with care transitions

CJR survey respondents first discharged to an IRF reported less satisfactory care transition experiences than control respondents. Roughly three fewer CJR respondents out of 100 indicated that they were discharged from the hospital at the right time (-3.0 percentage points, p<0.10), received the right amount of post-discharge care (-2.7 percentage points, p<0.10), or had all the necessary medical equipment needed when they went home (-3.3 percentage points, p<0.10) (Exhibit 37).

Caregiver help

Approximately 90% of CJR and control survey respondents first discharged to an IRF received help from a caregiver after returning home, and this proportion was not statistically significantly different between the two groups (Appendix H). Among those who received caregiver help at home, CJR respondents required more caregiver help at home than control respondents based on

---

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

three ADLs scored from 0 to 100. Differences in the composite across all measures was -4.9 points out of 100 (p<0.10, Exhibit 37). CJR respondents reported needing more help putting on or taking off clothes (-5.3 points, p<0.10), bathing (-5.4 points, p<0.05), and using the toilet (-3.3 points, p<0.10) than control respondents.

Exhibit 37:   Relatively fewer CJR survey respondents first discharged to an IRF reported satisfactory care transition experiences and CJR respondents required more caregiver help at home than control respondents

Domain	Survey measure	CJR	Control group	Difference between CJR and control groups
Satisfaction with overall recovery[a]	Satisfaction with overall recovery since leaving the hospital	76.4	78.2	-1.8
Satisfaction with care management[a]	Composite measure of satisfaction with care management	78.2	80.8	-2.6
	Healthcare providers listened to preferences	74.4	76.6	-2.1
	Satisfaction with discharge destination	78.8	78.4	0.4
	Satisfaction with care coordination	77.8	78.6	-0.8
	Satisfaction with treatment instructions	79.3	82.0	-2.7
Experience with care transitions[b]	Discharged from the hospital at the right time	84.4	87.4	-3.0
	Received the right amount of post-discharge care	83.7	86.4	-2.7
	Had all the medical equipment needed at home	90.7	94.0	-3.3
Caregiver help[c]	Composite measure of caregiver help	70.6	75.5	-4.9
	Help needed putting on or taking off clothes	63.0	68.3	-5.3
	Help needed bathing	64.7	70.1	-5.4
	Help needed using the toilet	82.3	85.6	-3.3

Source: CJR evaluation team analysis of patient survey data for patients first discharged to an IRF in March, April, August, or September 2018.

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

Survey sample sizes by setting are reported in Exhibit 30.

IRF = inpatient rehabilitation facility.

[a]   Measures in domain reported as satisfaction from 0 points (very dissatisfied) to 100 points (very satisfied).

[b]   Measures in domain reported as percent of respondents answering "yes" to a given question.

[c]   Measures in domain reported as amount of caregiver help needed from 0 points (complete help needed) to 100 points (no help needed).

Summary

CJR patients first discharged to an IRF were more complex during the intervention period than during the baseline period, based on observable demographic and clinical factors, compared to control patients. This is consistent with the shift toward less intensive PAC settings for the CJR group, which left the most complex patients in the IRF.

We found some evidence of worse longer term self-reported functional improvement among CJR survey respondents first discharged to an IRF. CJR survey respondents reported less improvement in some measures of functional status and pain than control respondents from before their surgery to after the episode. CJR and control patients, however, had similar improvements in functional status during their IRF stay, based on one assessment-based measure. Small differences between the survey and assessment results are most likely due to differences in the corresponding measures, not necessarily differences between patients' short-term and longer-term functional outcomes.

CJR and control respondents reported similar satisfaction with care management and with overall recovery, but fewer CJR respondents reported satisfactory care transition experiences and required more caregiver help at home compared to control respondents. The magnitude of the estimates, however, were small and may reflect some downward bias.

Patients first discharged to a SNF

Changes in patient mix

There was little evidence of an increase in the complexity of CJR patients first discharged to a SNF, relative to control patients. In the SNF setting, there was a relative increase in the proportion of CJR patients who were 80 years or older (p<0.10), as well as relative increases in CJR patients' average early-loss ADL scores (i.e., dressing and personal hygiene, p<0.10) and mid-loss ADL scores (i.e., transfer, locomotion on unit, and walking in corridor, p<0.10) measured at the time of admission to the SNF (Exhibit 38, Appendix I). For both ADL scores, the average CJR patient's functional status remained stable from baseline to intervention, while the average control patient's functional status improved. There were few differences in observable measures of complexity, although changes reflect increases in complexity that are consistent with the shift away from institutional PAC under the CJR model. To the extent that changes in discharge destination reflect indicators of patient complexity that cannot be observed in the data (e.g., lack of caregiver at home; comorbid conditions not captured by the HCC score), results will appear more negative for CJR patients.

Exhibit 38: Little evidence of an increase in patient complexity for CJR patients first discharged to a SNF, PY1-3

	Characteristic	Net difference	Lower episode payments	Higher episode payments	*p*- value
Demographics	Age, 80+	2.1			<0.10
	Sex, Female	-0.3			0.64
	Race, Black	-1.0			0.12
Social Determinants	Eligible for Medicaid	-0.9			0.45
	Disability, no ESRD	-0.6			0.35
Health status	HCC Score	0.1			0.22
	Obesity	-1.3			0.43
	Diabetes	0.9			0.21
	Hypertension	0.7			0.25
	Dementia	1.5			0.13
	Congestive Heart Failure	0.1			0.90
	Hip Fracture	1.9			0.39
	MS-DRG 469	0.8			0.29
	Bathing poor	0.9			0.67
	Cognition not intact	0.1			0.94
	Severe cognitive impairment	0.1			0.88
	Early-loss ADL score	0.1			<0.10
	Mid-loss ADL score	0.2			<0.10
	Late-loss ADL score	0.2			0.13
Prior use	ACH Stay	0.7			0.11
	HH Use	-0.3			0.79
	IRF Stay	0.0			0.92
	SNF Stay	0.1			0.60
	Any Prior Care	-0.1			0.91

Net difference (90% CI)

Source: CJR evaluation team analysis of Medicare claims, enrollment and 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 2018 (intervention).

Notes: Net differences at the 99%, 95%, or 90% significance level are indicated by red, orange, or yellow shaded cells, respectively, and are considered to indicate a relative increase or decrease in patient complexity.

ACH = acute care hospital, ADLs = activities of daily living, CI = confidence interval, ESRD = end-stage renal disease, HCC = hierarchical condition category, HH = home health, IRF = inpatient rehabilitation facility, MDS = Minimum Data Set, MS-DRG = Medicare Severity-Diagnosis Related Group, PY = performance year, SNF = skilled nursing facility.

Functional status and pain

Patient survey results

Overall, CJR and control survey respondents first discharged to a SNF experienced similar improvement in all eight measures of functional status and pain from before their surgery to the time of the survey (Exhibit 39). CJR respondents reported 3.7% more improvement than control

respondents in walking by themselves without rest, (p<0.10). No other differences between the two groups were statistically significant.

Exhibit 39:   CJR and control survey respondents first discharged to a SNF generally reported similar changes in functional status

Survey measure	Response range	Mean change in self-reported measure from before the hospitalization to after the episode *Higher value represents a more favorable change*		Difference between CJR and control groups (% difference)[a]
		CJR	Control group	
Ability to walk by yourself without resting	-4 to 4	0.56	0.47	0.09 (3.7%)
Difficulty walking up or down 12 stairs	-3 to 3	0.66	0.63	0.03 (1.5%)
Difficulty rising from sitting	-4 to 4	1.10	1.11	-0.01 (-0.3%)
Difficulty standing	-4 to 4	1.04	1.02	0.01 (0.5%)
Use of a mobility aid	-2 to 2	0.03	0.03	0.01 (0.4%)
Difficulty getting on/off the toilet	-4 to 4	1.30	1.31	-0.02 (-0.5%)
Frequency that pain interferes with normal activities	-4 to 4	1.89	1.91	-0.02 (-0.8%)
Medication use for pain in the joint you had replaced	-3 to 3	0.57	0.57	-0.00 (-0.1%)

Source: CJR evaluation team analysis of patient survey data for patients first discharged to a SNF in March, April, August, or September 2018.

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

Survey sample sizes by setting are reported in Exhibit 30.

SNF = skilled nursing facility.

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

Assessment-based results

The assessment-based results are mixed for patients first discharged to a SNF. The proportion of CJR patients who improved in toilet use during their SNF stay decreased relative to the control group. 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 3.0 percentage points from baseline to intervention (from 44.0% to 41.0%); for the control group it increased by 0.7 percentage points (from 46.9% to 47.6%) (Exhibit 40). As a result, the proportion of CJR patients who improved in toilet use decreased by 3.7 percentage points relative to the control group, or 8.4% from the CJR baseline (p<0.10, Appendix B, Exhibit B-2).

Exhibit 40:   For patients first discharged to a SNF, the proportion with improved toilet use decreased from baseline to intervention for CJR patients and increased for control patients, PY1-3
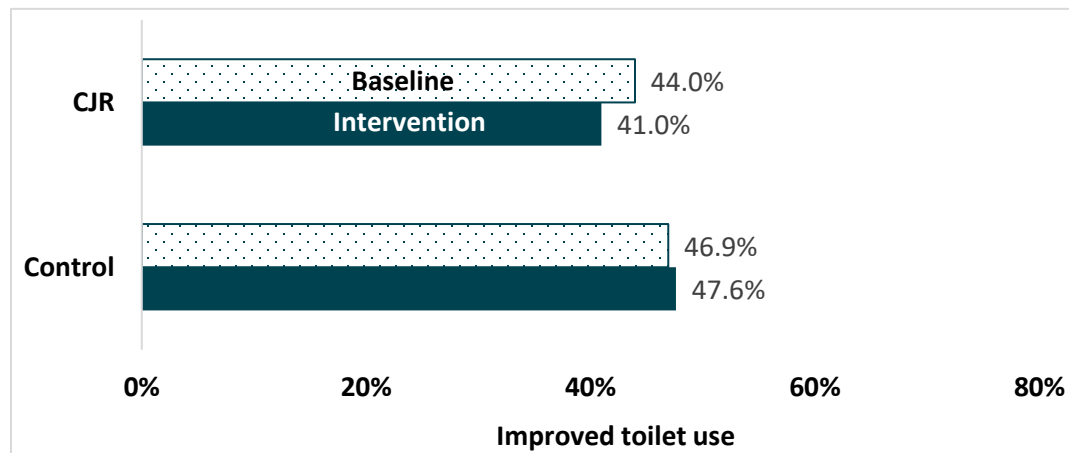
Source: CJR evaluation team analysis of MDS data for episodes initiated in 2012 through 2014 that ended between April 2012 and March 2015 (baseline) and episodes initiated during or after April 2016 that ended by September 2018 (intervention).

Note: MDS = Minimum Data Set, PY = performance year, SNF = skilled nursing facility.

The proportion of CJR patients without pain increased by 16.6 percentage points from baseline to intervention (from 58.2% to 74.8%). For the control group, the proportion of patients without pain increased by 13.4 percentage points (from 53.0% to 66.4%) (Exhibit 41). As a result, the proportion of CJR patients without pain increased by 3.2 percentage points relative to the control patients, or 5.6% from the CJR baseline (p<0.05, Appendix B, Exhibit B-2).

Exhibit 41:   For patients first discharged to a SNF, the proportion without moderate to severe pain increased more for CJR than for control patients, PY1-3
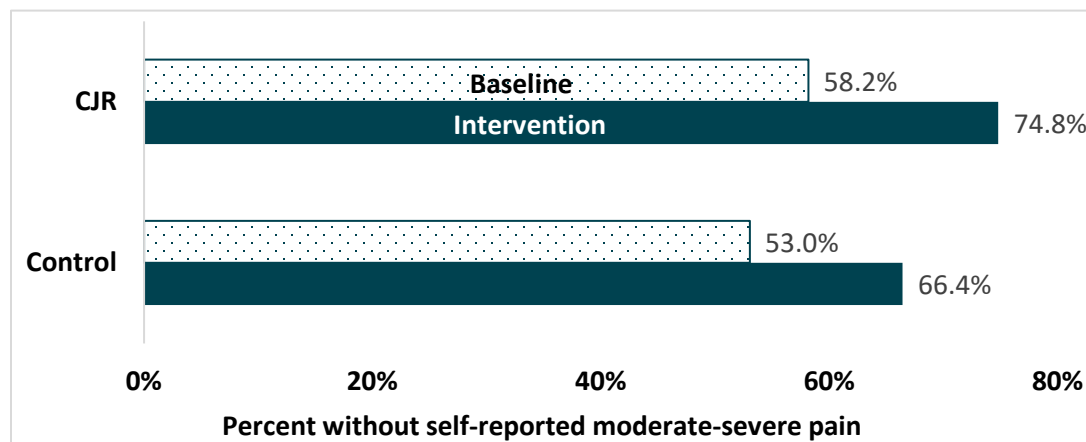
Source: CJR evaluation team analysis of MDS data for episodes initiated in 2012 through 2014 that ended between April 2012 and March 2015 (baseline) and episodes initiated during or after April 2016 that ended by September 2018 (intervention).

Notes: These results did not fail the restrictive parallel trends test, but do fail the linear parallel trends test at the 5% level, and therefore need to be interpreted with caution.

MDS = Minimum Data Set, PY = performance year, SNF = skilled nursing facility.

Between the baseline and intervention periods, changes in the proportion of patients whose mobility improved during their SNF stay were not statistically different between the CJR and control groups (Appendix B, Exhibit B-2).

Sensitivity findings

We conducted additional analyses to better understand the changes in functional status of CJR patients first discharged to a SNF. Consistent with a reduction in the number of SNF days (Section II.B.3), the number of days between patient assessments (upon SNF admission and discharge) decreased by 1.8 days more for CJR patients than for control patients ($p<0.05$, Appendix J). This indicates that CJR patients had less time to improve in functional status while in the SNF, which could have contributed to the relative decrease in the proportion of CJR patients with improved toilet use. After controlling for the number of days between entry and exit assessments in the risk adjustment model, the estimated DiD impact of the CJR model on toilet use decreased from 8.4% to 5.4% of the CJR baseline and was no longer statistically significant (Appendix J).[50] Thus, for CJR patients first discharged to a SNF, their shorter length of stay explains some of the relative decrease in the proportion whose toilet use improved while in the SNF.

Satisfaction, experience with care transitions, and caregiver help[51]

Satisfaction

There were no statistically significant differences between CJR and control respondents on any measure of satisfaction related to overall recovery or care management.

Experience with care transitions

CJR survey respondents first discharged to a SNF reported less satisfactory care transition experiences than control respondents. Compared to control respondents, approximately three fewer CJR respondents out of every 100 reported being discharged from the hospital at the right time (-3.5 percentage points, $p<0.01$) and receiving the right amount of post-discharge care (-2.6 percentage points, $p<0.05$, Exhibit 42). There was no statistically significant difference in the proportion of respondents who indicated that they received all the medical equipment they needed when they went home.

Caregiver help

Almost 90% of CJR and control survey respondents first discharged to a SNF received help from a caregiver after returning home, and this proportion was not statistically significantly different between the two groups (Appendix H). Among those receiving caregiver help when they returned home, CJR respondents required more caregiver help than control respondents. CJR respondents

---

[50] 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 proportion of CJR patients whose functional status improved can be explained by its association with the relative change in number of days between assessments among CJR patients.

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

scored 2.6 points lower out of 100 on a composite measure of three ADLs (p<0.10, Exhibit 42). CJR respondents reported needing more help than control respondents with putting on or taking off clothes (-2.9 points, p<0.10) and bathing (-3.6 points, p<0.05).

Exhibit 42:   Fewer CJR survey respondents first discharged to a SNF reported satisfactory care transition experiences and CJR respondents required more caregiver help at home than control respondents

Domain	Survey measure	CJR	Control group	Difference between CJR and control groups
Satisfaction with overall recovery[a]	Satisfaction with overall recovery since leaving the hospital	77.0	78.5	-1.5
Satisfaction with care management[a]	Composite measure of satisfaction with care management	76.2	77.0	-0.8
	Healthcare providers listened to preferences	72.8	74.4	-1.6
	Satisfaction with discharge destination	70.9	71.0	-0.1
	Satisfaction with care coordination	76.9	77.9	-1.0
	Satisfaction with treatment instructions	80.3	81.1	-0.8
Experience with care transitions[b]	Discharged from the hospital at the right time	83.6	87.0	-3.5
	Received the right amount of post-discharge care	81.4	84.0	-2.6
	Had all the medical equipment needed at home	91.0	91.6	-0.6
Caregiver help[c]	Composite measure of caregiver help	72.8	75.4	-2.6
	Help needed putting on or taking off clothes	66.7	69.5	-2.9
	Help needed bathing	67.4	71.0	-3.6
	Help needed using the toilet	82.3	83.4	-1.1

Source: CJR evaluation team analysis of patient survey data for patients first discharged to a SNF in March, April, August, or September 2018.

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

Survey sample sizes by setting are reported in Exhibit 30.

SNF = skilled nursing facility.

[a]   Measures in domain reported as satisfaction from 0 points (very dissatisfied) to 100 points (very satisfied).

[b]   Measures in domain reported as percent of respondents answering "yes" to a given question.

[c]   Measures in domain reported as amount of caregiver help needed from 0 points (complete help needed) to 100 points (no help needed).

Summary

There was little evidence that CJR patients first discharged to a SNF were more complex during the intervention period than the baseline period based on observable factors. CJR and control survey respondents reported similar improvement in functional status and pain from before their surgery to the time of the survey. Measures of functional status and pain from the assessment data were mixed. There was a relative decrease in the proportion of CJR patients with improved toilet use and a relative increase in the proportion of CJR patients without pain during their SNF stay.

CJR and control survey respondents reported similar satisfaction with care management and with overall recovery, but CJR respondents reported less satisfactory care transition experiences and required more caregiver help at home than control respondents. These differences were limited and could be due to differing patient complexity in the SNF setting between the CJR and control group that cannot be observed in the data (e.g., lack of caregiver at home; comorbid conditions not captured by the HCC score).

Patients first discharged to an HHA

Changes in patient mix

Changes in patient characteristics between baseline and intervention were generally similar for CJR and control group patients first discharged to an HHA (Exhibit 43). However, there were relative increases in the proportion of CJR patients who were female ($p<0.10$), had diabetes ($p<0.01$), had hypertension ($p<0.05$), or had difficulty bathing ($p<0.10$). The proportion of patients with a given characteristic increased among CJR patients for all measures, and among control patients for all measures except diabetes, so differential changes in patient complexity between the CJR and control group were due to a greater increase in the probability of a given characteristic in the CJR group.

Exhibit 43: Little evidence of a relative change in patient complexity for CJR patients first discharged to an HHA, PY1-3

Characteristic	Net difference	Lower episode payments	Higher episode payments	*p*- value
Demographics				
Age, 80+	0.5			0.28
Sex, Female	1.1			<0.10
Race, Black	-0.1			0.82
Social Determinants				
Eligible for Medicaid	-0.1			0.88
Disability, no ESRD	-0.2			0.80
Health status				
HCC Score	0.0			0.68
Obesity	0.7			0.78
Diabetes	2.4			<0.01
Hypertension	1.6			<0.05
Dementia	-0.1			0.74
Congestive Heart Failure	0.5			0.23
Hip Fracture	-0.1			0.85
MS-DRG 469	0.0			0.82
Toilet transferring	-2.4			0.38
Transferring	-1.5			0.63
Ambulation/locomotion	1.9			0.44
Lower body dressing	0.0			0.96
Upper body dressing	2.5			0.23
Bathing	3.4			<0.10
Toileting hygiene	0.7			0.80
Grooming	0.9			0.71
Cognitive functioning	0.0			0.91
Confusion	0.0			0.88
Memory deficit	-0.1			0.84
Impaired decision-making	-0.5			0.50
Overall status	1.8			0.24
Prior use				
ACH Stay	-0.3			0.45
HH Use	-0.8			0.12
IRF Stay	0.0			0.49
SNF Stay	-0.1			0.14
Any Prior Care	-0.8			0.46

Net difference (90% CI)

Source: CJR evaluation team analysis of Medicare claims, enrollment and OASIS data for episodes initiated in 2012 through 2014 that ended between April 2012 and March 2015 (baseline) and episodes initiated during or after April 2016 that ended by December 2018 (intervention).

Notes: Net differences at the 99%, 95%, or 90% significance level are indicated by red, orange, or yellow shaded cells, respectively, and are considered to indicate a relative increase or decrease in patient complexity.

ACH = acute care hospital, CI = confidence interval, ESRD = end-stage renal disease, HCC = hierarchical condition category, HH = home health, HHA = home health agency, IRF = inpatient rehabilitation facility, MS-DRG = Medicare Severity-Diagnosis Related Group, OASIS = Outcome and Assessment Information Set, PY = performance year, SNF = skilled nursing facility.

Functional status and pain

Patient survey results

On average, CJR and control respondents first discharged to an HHA reported improvement in functional status and pain, from before their surgery to the time of the survey, but CJR respondents experienced less improvement than control respondents in two of eight measures (Exhibit 44). CJR respondents reported less improvement in rising from sitting ($p<0.05$) and increased dependence on a mobility device ($p<0.05$), although both differences were less than 0.05 or 2% of the average CJR recalled status prior to the hospitalization. There were no statistically significant differences between the two groups on the other six measures.

Exhibit 44:   CJR survey respondents first discharged to an HHA reported less improvement in rising from sitting and more dependence on a mobility device than control respondents

Survey measure	Response range	Mean change in self-reported measure from before the hospitalization to after the episode *Higher value represents a more favorable change*		Difference between CJR and control groups (% difference)[a]
		CJR	**Control group**	
Ability to walk by yourself without resting	-4 to 4	0.89	0.88	0.01 (0.5%)
Difficulty walking up or down 12 stairs	-3 to 3	0.84	0.86	-0.02 (-1.0%)
Difficulty rising from sitting	-4 to 4	1.28	1.32	-0.04 (-1.5%)
Difficulty standing	-4 to 4	1.26	1.28	-0.02 (-0.9%)
Use of a mobility aid	-2 to 2	0.25	0.28	-0.03 (-1.3%)
Difficulty getting on/off the toilet	-4 to 4	1.45	1.46	-0.01 (-0.5%)
Frequency that pain interferes with normal activities	-4 to 4	1.99	2.00	-0.01 (-0.7%)
Medication use for pain in the joint you had replaced	-3 to 3	0.62	0.64	-0.03 (-1.0%)

Source: CJR evaluation team analysis of patient survey data for patients first discharged to an HHA in March, April, August, or September 2018.

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

Survey sample sizes by setting are reported in Exhibit 30.

HHA = home health agency.

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

Assessment-based results

There were no statistically significant differences in shorter-term measures of functional status or pain between CJR and control patients first discharged to an HHA from baseline to intervention. In both the baseline and intervention periods, at least three-quarters of CJR and control patients improved in bed transferring, improved in ambulation/locomotion, or had reduced pain while receiving HH care (Appendix B, Exhibit B-2).

Satisfaction, experience with care transitions, and caregiver help[52]

Satisfaction

CJR respondents first discharged to an HHA were less satisfied with their overall recovery than control respondents (-1.1 points out of 100, $p < 0.05$, Exhibit 45). CJR respondents were also less satisfied than control respondents with the extent to which health care providers listened to their preferences (-1.6 points, $p < 0.05$) and with coordination of their care (-1.0 points, $p < 0.10$). Differences between the two groups for the measures of satisfaction with care management and the summary composite measure were not statistically significant.

Experience with care transitions

Among patients first discharged to an HHA, similar percentages of CJR and control respondents indicated that they were discharged at the right time, received the right amount of post-discharge care, and had all the medical equipment needed at home.

Caregiver help

Approximately 97% of CJR and control survey respondents first discharged to an HHA received help from a caregiver after returning home, and this proportion was not statistically significantly different between the two groups (Appendix H). CJR and control respondents required similar levels of caregiver help at home based on a composite measure of three ADLs. However, CJR respondents needed more help bathing than did control respondents (-1.9 points out of 100, $p < 0.10$, Exhibit 45), but there were no statistically significant differences between the two groups in help needed putting on or take off clothes or using the toilet.

---

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

Exhibit 45:   CJR survey respondents first discharged to an HHA were less satisfied with their overall recovery and with some aspects of care management than control respondents

Domain	Survey measure	CJR	Control group	Difference between CJR and control groups
Satisfaction with overall recovery[a]	Satisfaction with overall recovery since leaving the hospital	81.2	82.4	-1.1
Satisfaction with care management[a]	Composite measure of satisfaction with care management	84.3	85.1	-0.8
	Healthcare providers listened to preferences	78.5	80.1	-1.6
	Satisfaction with discharge destination	85.9	86.1	-0.2
	Satisfaction with care coordination	83.9	84.8	-1.0
	Satisfaction with treatment instructions	85.3	86.2	-0.9
Experience with care transitions[b]	Discharged from the hospital at the right time	88.1	88.5	-0.4
	Received the right amount of post-discharge care	86.8	87.2	-0.4
	Had all the medical equipment needed at home	91.6	92.2	-0.6
Caregiver help[c]	Composite measure of caregiver help	66.1	67.4	-1.3
	Help needed putting on or taking off clothes	57.7	59.2	-1.4
	Help needed bathing	63.3	65.2	-1.9
	Help needed using the toilet	78.5	80.1	-1.6

Source: CJR evaluation team analysis of patient survey data for patients first discharged to an HHA in March, April, August, or September 2018.

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

Survey sample sizes by setting are reported in Exhibit 30.

HHA = home health agency.

[a]   Measures in domain reported as satisfaction from 0 points (very dissatisfied) to 100 points (very satisfied).

[b]   Measures in domain reported as percent of respondents answering "yes" to a given question.

[c]   Measures in domain reported as amount of caregiver help needed from 0 points (complete help needed) to 100 points (no help needed).

Summary

There was little evidence that CJR patients first discharged to an HHA were more complex during the intervention period than the baseline period, at least for factors that can be observed in the data. CJR patients reported less improvement in functional status than control respondents from before their surgery to the time of the survey for two of eight measures, and differences were less than 2 percent in magnitude. The HHA clinical assessments indicate that the two groups had similar improvements in functional status and pain while receiving HH care. Small differences between the survey and assessment results are most likely due to differences in the corresponding measures, not necessarily differences between patients' short-term and longer-term functional outcomes.

CJR survey respondents were less satisfied with their overall recovery and with some aspects of care management than control respondents. These differences were limited and could be driven by

unobserved differences in patient complexity between the CJR and the control group in the HHA setting (e.g., lack of caregiver at home; comorbid conditions not captured by the HCC scores). The two groups reported similar satisfaction with care transition experiences and required similar levels of caregiver help at home.

Patients discharged home without HH care

Changes in patient mix

We found no consistent evidence that CJR patients discharged directly home without HH care were more or less complex in the intervention period than the baseline, relative to control patients, which is consistent with no differential change in the proportion of patients discharged home with HH care (Appendix I).

Functional status and pain

Patient survey results

CJR and control survey respondents who were discharged directly home without HH care reported similar improvement in all eight measures of functional status and pain, from before their surgery to the time of the survey (Exhibit 46). No difference between the two groups was statistically significant.

Exhibit 46:   CJR and control respondents discharged home without HH care reported similar improvement in functional status and pain

Survey measure	Response range	Mean change in self-reported measure from before the hospitalization to after the episode *Higher value represents a more favorable change*		Difference between CJR and control groups (% difference)[a]
		CJR	Control group	
Ability to walk by yourself without resting	-4 to 4	0.90	0.85	0.05 (1.7%)
Difficulty walking up or down 12 stairs	-3 to 3	0.82	0.82	0.00 (0.2%)
Difficulty rising from sitting	-4 to 4	1.26	1.25	0.01 (0.5%)
Difficulty standing	-4 to 4	1.21	1.19	0.02 (0.7%)
Use of a mobility aid	-2 to 2	0.25	0.25	0.00 (0.1%)
Difficulty getting on/off the toilet	-4 to 4	1.41	1.39	0.02 (0.7%)
Frequency that pain interferes with normal activities	-4 to 4	2.00	2.02	-0.02 (-1.1%)
Medication use for pain in the joint you had replaced	-3 to 3	0.65	0.60	0.05 (1.7%)

Source: CJR evaluation team analysis of patient survey data for patients discharged home without HH care in March, April, August, or September 2018.

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

Survey sample sizes by setting are reported in Exhibit 30.

HH = home health.

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

Satisfaction, experience with care transitions, and caregiver help[53]

Satisfaction

CJR respondents discharged home without HH were more satisfied with care management than control respondents, based on a composite of four satisfaction measures (1.9 points out of 100, p<0.10, Exhibit 47). The difference between the groups in the composite score was primarily driven by satisfaction with discharge destination: CJR respondents were 2.3 points more satisfied with their discharge destination (out of 100 points) than were control respondents when they were discharged home without HH (p<0.10). Differences between the groups on the other three measures of satisfaction with care management and satisfaction with overall recovery measure were not statistically significant.

Experience with care transitions

Among survey respondents discharged directly home without HH care, similar percentages of CJR and control respondents indicated that they were discharged at the right time, received the right amount of post-discharge care, and had all the medical equipment needed at home.

Caregiver help

Approximately 97% of CJR and control survey respondents discharged directly home without HH care received at least some help from a caregiver after returning home, and this proportion was not statistically significantly different between the two groups (Appendix H). The two groups were also similar in the level of help they needed from caregivers at home on all three ADLs (Exhibit 47).

---

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

Exhibit 47:   CJR respondents discharged home without HH care were more satisfied with some aspects of care management than control respondents

Domain	Survey measure	CJR	Control group	Difference between CJR and control groups
Satisfaction with overall recovery[a]	Satisfaction with overall recovery since leaving the hospital	81.1	79.3	1.8
Satisfaction with care management[a]	Composite measure of satisfaction with care management	84.4	82.6	1.9
	Healthcare providers listened to preferences	78.7	77.7	1.0
	Satisfaction with discharge destination	85.8	83.5	2.3
	Satisfaction with care coordination	83.6	82.0	1.6
	Satisfaction with treatment instructions	84.3	82.6	1.7
Experience with care transitions[b]	Discharged from the hospital at the right time	87.0	86.8	0.2
	Received the right amount of post-discharge care	84.3	84.8	-0.5
	Had all the medical equipment needed at home	93.6	92.4	1.2
Caregiver help[c]	Composite measure of caregiver help	67.2	67.1	0.1
	Help needed putting on or taking off clothes	58.6	57.9	0.8
	Help needed bathing	66.1	64.9	1.3
	Help needed using the toilet	79.9	77.8	2.1

Source: CJR evaluation team analysis of patient survey data for patients discharged home without HH care in March, April, August, or September 2018.

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

Survey sample sizes by setting are reported in Exhibit 30.

HH = home health.

[a]   Measures in domain reported as satisfaction from 0 (very dissatisfied) to 100 (very satisfied).

[b]   Measures in domain reported as percent of respondents answering "yes" to a given question.

[c]   Measures in domain reported as amount of caregiver help needed from 0 (complete help needed) to 100 (no help needed).

Summary

There was no differential change in the percent of patients first discharged home without HH care from baseline to intervention between CJR and control groups. Changes in patient complexity between baseline and intervention were similar for CJR and control patients discharged home without HH care. CJR and control survey respondents reported similar improvement in functional status and pain from before their surgery to the time of the survey. CJR respondents were more satisfied with some aspects of care management than control respondents, but this difference was limited and both groups reported similar satisfaction with overall recovery and care transition experiences and required similar levels of caregiver help at home.

d.   Conclusion

The self-reported outcomes from the patient survey indicate that overall CJR and control respondents had similar improvement in functional status and pain from the week before their

surgery to after the end of the episode. On average, across all respondents, longer-term functional status and pain outcomes were similar for the CJR and control groups, despite the shift away from institutional PAC under the CJR model.

There was some variation in longer-term functional recovery for patients discharged to different PAC settings, which may be due to the shift away from institutional PAC during the CJR model. CJR and control respondents first discharged to SNFs and home without HH had similar improvement in functional status and pain. We found some evidence of less functional improvement among CJR survey respondents first discharged to an IRF, and to a lesser extent among patients first discharged to an HHA.

We examined shorter-term changes in functional status and pain for patients first discharged to an IRF, SNF, or HHA using PAC assessment data. With the exception of patients first discharged to a SNF, we found that CJR and control patients had similar improvements in functional status and pain during their first PAC stay. The results were mixed for patients who were first discharged to a SNF. There was a relative decrease in the proportion of CJR patients who improved in toilet use during their SNF stay. At the same time, the proportion of SNF patients without moderate to severe pain increased, signifying an improvement in quality of care. CJR patients discharged to a SNF had less time to improve their functional status than control patients because of their shorter SNF stays. This explained some, but not all, of the relative decrease in the proportion of patients with improvement in toilet use.

In aggregate, CJR and control survey respondents reported similar satisfaction with overall recovery and with care management. There were, however, some differences between CJR and control respondents in care transition experiences. CJR respondents were slightly less likely than control respondents to indicate that they were discharged from the hospital at the right time or received the right amount of care after discharge. CJR respondents also indicated that they needed a little more caregiver help at home than control respondents. Aggregate differences in care transition experiences and caregiver help were driven by CJR respondents discharged to an IRF or SNF.

Analysis of outcomes within settings provides reassurance that the aggregate results were not masking large variations in outcomes, as differences between the CJR and control groups within each setting were minimal. The survey analysis did suggest some differences between CJR and control outcomes that were not seen in the assessment data analysis. It is difficult, however, to compare results derived from these different data sources, especially when impacts are small in magnitude. The survey and assessment results are based on different measures, so small differences in results may be due to differences in measures rather than differences in short-term and longer-term patient outcomes.

Furthermore, because the CJR model caused shifting of patients from more intense to less intense PAC settings, estimates by PAC setting from both survey and assessment data may be

biased downward (appearing to indicate worse CJR outcomes) due to shifts in patient complexity in each setting that we are unable to measure.

6. 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, 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.

a. Key Findings

- In the first three performance years, the CJR model did not have a statistically significant impact on the volume of elective LEJR discharges (inpatient plus outpatient) in the CJR markets that were mandatory as of 2018.

- For elective MS-DRG 470, the least complex episode group, the CJR patient population was relatively healthier in the intervention period than in the baseline period, which could make it easier for CJR participant hospitals to achieve payment and quality targets and thus receive reconciliation payments.

- For the more complex episode types – elective MS-DRG 469, fracture MS-DRG 470, and fracture MS-DRG 469 – the CJR patient population was not relatively healthier in the intervention period.

- We cannot draw a conclusion about the impact of the CJR model on payments for services provided during the 30 days following the episode.

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

The mandatory CJR model targets LEJR surgery in part because of its prevalence in the Medicare population, with more than 400,000 procedures in 2014 and growth projected to continue.[54] The volume of these procedures has been trending upward since the 1990s, with rates of total hip and

---

[54] 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.

knee replacements approximately doubling among those 45 and older between 2000 and 2010.[55,56] Exhibit 48 shows an increase in the elective LEJR discharge rate for the Medicare FFS population from 2007.

Exhibit 48:   Increasing national trend in the elective LEJR discharge rate since 2007
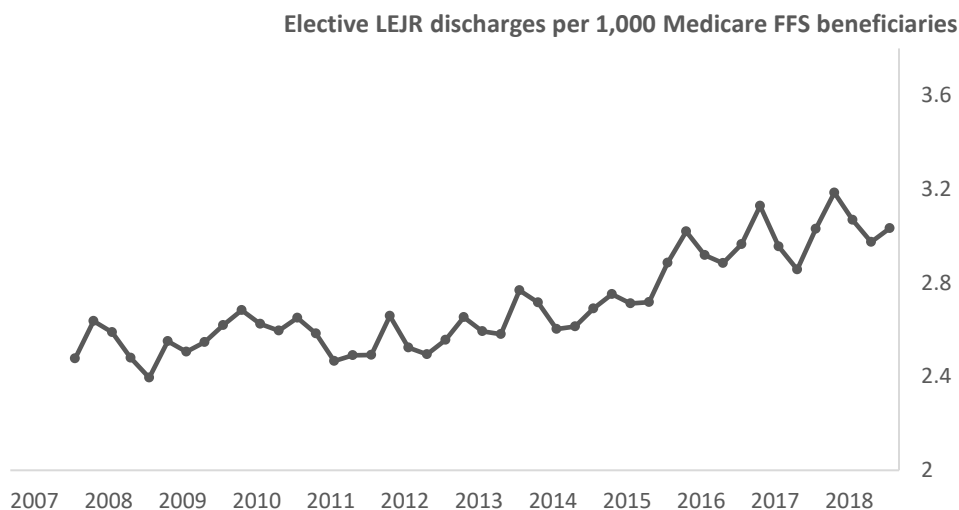
Elective LEJR discharges per 1,000 Medicare FFS beneficiaries

Source: CJR evaluation team analysis of Medicare claims and enrollment data for discharges from October 2007 through December 2018.

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

In the context of overall growth in 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. Participant hospitals may be able to reduce average episode payments by providing elective 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 due to the CJR model would be offset by the payments for these additional episodes.

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

---

[55]   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.

[56]   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.

agreements with referring physicians. Shifts in volume across providers, however, would likely not have much effect on Medicare savings due to the CJR model.[57]

a. Methods

We analyzed the impact of the CJR model on market-level volume of elective LEJR discharges because sampling for participation in the model occurred at the MSA level and all acute care hospitals paid under the Medicare IPPS in the MSA, with few exceptions, were required to participate. This analysis included all elective LEJR discharges in the 34 mandatory CJR MSAs relative to control group MSAs, weighted based on the probability that the MSAs were selected into the mandatory group.[58] Elective LEJR discharges included: 1) LEJR procedures performed as hospital inpatient procedures (discharged under MS-DRG 469 or 470 and without fracture diagnosis codes); and 2) starting in 2018, total knee arthroplasty (TKA) procedures performed as hospital outpatient procedures (Current Procedural Terminology (CPT) code 27447). TKA was removed from the inpatient only list starting January 1, 2018. We included outpatient TKA in the elective discharge rate to evaluate the extent to which the CJR model resulted in increases in overall volume, regardless of whether the TKA was performed as an inpatient or outpatient procedure. See Section II.A.2 of the report for additional discussion about outpatient TKA.

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 (BPCI) initiative had an LEJR episode that was structured similarly to CJR episodes, and therefore had similar incentives for boosting volume of elective LEJR surgeries. We also controlled for hospital participation in BPCI by including a BPCI "dose" for each MSA.

The dose variables for CJR and the BPCI initiative were defined as the baseline market share of providers that have ever been in the CJR model or the risk-bearing phase of the BPCI initiative, respectively. The baseline period for dose measurement was from October 2009 through September 2012, the three years before the first BPCI intervention period.[59]

We defined three CJR intervention periods:

■ **Interim** begins the quarter that the CJR model was announced (July 2015) and ends the quarter before it was implemented (March 2016).[60]

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

---

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

[58] All elective LEJR discharges occurring in a MSA were included, whether they occurred at a CJR hospital or non-participating hospital.

[59] 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.

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

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

We interacted each of the intervention period variables with the CJR dose and with the BPCI dose. We also controlled for market-level characteristics, market and quarterly fixed effects, and a market-specific linear time trend. We report differences in discharge rates between CJR and control group MSAs for the three different intervention periods and cumulatively across performance years (PY1-PY3). We also tested whether the results were significantly different across intervention periods. See Appendix E (Section VII) for a full description of methodology.

b.   Results

There was no statistically significant difference in the volume of elective LEJR discharges between mandatory CJR MSAs and control group MSAs. The estimated change in the elective LEJR discharge rate in the first three performance years due to the CJR model was a decrease of 0.06 per 1,000 Medicare FFS beneficiaries (p=0.31, Exhibit 49). This result was stable over time with similar differences in discharge rates between CJR and control group MSAs in the interim period (-0.04), PY1-2 (-0.06), and PY3 (-0.05). The differences in rates for each intervention period are not statistically significantly different from one another (p=0.61). There is limited ability for hospitals to induce volume for fracture episodes, and as expected, we observed no change in fracture discharge rates due to the CJR model.[61]

Exhibit 49:   The CJR model did not have a statistically significant impact on the volume of elective LEJR discharges in mandatory MSAs

Period	Predicted CJR MSA discharge rate (per 1,000 FFS beneficiaries)	Predicted control group MSA discharge rate (per 1,000 FFS beneficiaries)	Difference in discharge rates (per 1,000 FFS beneficiaries)	p-value
Interim	2.82	2.86	-0.04	0.32
PY1-2	2.91	2.97	-0.06	0.24
PY3	3.02	3.07	-0.05	0.49
PY1-3	2.95	3.01	-0.06	0.31

Source: CJR evaluation team analysis of Medicare claims and enrollment data for discharges from October 2007 through December 2018

Notes: R-squared = 0.93.

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

Interim includes discharges from July 2016 to March 2016. PY1-PY2 includes discharges from April 2016 through December 2017. PY3 includes discharges from January 2018 through December 2018.

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

---

[61]   The estimated change in the fracture LEJR discharge rate in the first three performance years due to the CJR model was a decrease of 0.01 per 1,000 FFS Medicare beneficiaries (p=0.63).

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

Less complex patients may require fewer resources and, therefore, have less costly episodes under the CJR model. Further, it may be easier to minimize negative quality indicators with a less complex patient population, resulting in higher quality scores. Indeed, CJR participant hospitals with less complex patient populations were more likely to have average episode payments below their quality-adjusted target prices and earn reconciliation payments (see Section II.C.1 for a discussion of the relationship between patient complexity and reconciliation payments). Certain patient characteristics—for example, being age 80 or older, being eligible for Medicaid, or having dementia—are associated with having higher episode payments. Throughout this annual report, we define patient complexity in terms of patient characteristics that are associated with higher episode payments.[62]

A potential unintended consequence of models such as the CJR model would be if participants receive reconciliation payments by shifting to a less complex mix of patients rather than through care redesign. Patient selection, whether intentional or not, would circumvent the goal of the model to reduce episode payments, and it could also reduce access to care for those patients who require more resources to treat. The CJR model's design was intended to limit patient selection. First, all hospitals in selected MSAs are participating in the model, which limits the ability to selectively admit the less complex patients. Second, target prices differ by four episode groups determined by fracture status (elective or fracture) and MS-DRG (470 or 469) (Exhibit 50), which is intended to account for differences in patient resource needs. As a result, the complexity of patients within an episode group would need to change to affect payment or quality outcomes.

---

[62] As discussed in the original final rule that established the CJR model, this definition of complexity is in line with how CMS pays Medicare Advantage plans; payments are tied to HCCs such that Medicare Advantage plans with relatively healthier patient populations are paid less than plans with average patient populations.

Exhibit 50:   There are quality-adjusted target prices for four episode groups, indicated by presence or absence of a fracture and MS-DRG

MS-DRG 470 elective	MS-DRG 469 elective	MS-DRG 470 fracture	MS-DRG 469 fracture
73% of CJR episodes	3% of CJR episodes	20% of CJR episodes	5% of CJR episodes
$23,279 Average quality-adjusted target price	$40,452 Average quality-adjusted target price	$43,610 Average quality-adjusted target price	$58,121 Average quality-adjusted target price

Source: CJR evaluation team analysis of Medicare claims and enrollment data for episodes initiated during or after April 2016 that ended by December 2018 (intervention) and analysis of CJR payment contractor target price data or PY1 through PY3.

*Notes***:**   Reported shares and means are averages over the three performance years for mandatory CJR hospitals.

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

There are several ways that patient complexity within an episode group could change. The complexity of patients within an episode group could change if a hospital upcoded the most complex patients in MS-DRG 470 (without major complications and comorbidities) to MS-DRG 469 (with major complications and comorbidities). If the hospital more completely documented conditions that could qualify as major complications or comorbidities to increase the number of patients coded as MS-DRG 469 instead of 470, then the complexity of patients in both groups would decrease, but at the same time, the target prices would remain the same. As a result, it would be easier for the hospital to have average episode payments in both groups that were below the target price. Another way that patient complexity could change is if healthier beneficiaries were encouraged to have LEJRs, which reduce the complexity of patients receiving LEJR while increasing the volume of LEJR procedures and thus Medicare spending. A hospital's patient complexity could also decrease if more complex Medicare beneficiaries were discouraged from having LEJR procedures.

Additionally, as discussed in Section II.A.2, the removal of TKA from the inpatient only list, effective January 2018, could also change patient complexity for CJR and control group hospitals. In theory, patients who could appropriately receive TKAs in the hospital outpatient setting would be less complex (i.e., have lower episode payments) than those in the inpatient setting. For instance, patients who could appropriately receive TKAs as outpatients would be expected to safely recover at home and forgo a SNF stay. It follows that as a hospital provides more outpatient TKAs, its remaining inpatient LEJR population would become more complex. Because mandatory CJR hospitals provided a lower share of TKAs in the outpatient setting than control group hospitals

(about 19% vs. 29%), the complexity of the CJR inpatient LEJR population would be expected to decrease, relative to the control group. Although we employ a rigorous risk adjustment methodology, it may not be robust enough to fully account for any differential change in patient complexity between the CJR and control group from baseline to intervention, which could bias our DiD estimates.

a. Methods

We analyzed changes in the mix of patients at mandatory CJR hospitals in two ways. First, we created a composite measure of patient complexity by predicting LEJR payments as a function of patient characteristics for each hospital. (Appendix E Section IX.B.2 details our calculation of predicted payments.) This has several advantages over testing for changes in multiple characteristics associated with episode payments. First, a composite measure reduces the risk of finding statistically significant changes that are due to chance. Second, because the measure summarizes changes in multiple characteristics, it avoids potentially inconclusive results that could occur if, for instance, changes for some characteristics were associated with lower episode payments and changes for other characteristics were associated with higher episode payments.

Using hospital-level averages of predicted payments, we computed the following ratio:

$$ratio_k = \frac{Predicted\ Y_{k,1}}{Predicted\ Y_{k,0}}$$

Where for each hospital *k*, *Predicted $Y_{k,0}$* is the average predicted payments during the baseline and *Predicted $Y_{k,1}$* is the average predicted payments during the intervention period. If this ratio of predicted payment is greater than 1, we conclude that the hospital's average patient complexity increased. If the ratio is less than 1, we conclude the hospital's average patient complexity decreased.

We also evaluated changes in discrete patient characteristics within each of the four episode groups to determine if the CJR patient population was different in the intervention period than it was 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. This complementary analysis provides insight into whether a change in the patient population is due to a consistent shift or if the patient population is becoming more complex along certain dimensions and less complex along others.

b. Results

Analyses of composite measure of patient complexity. As measured by the predicted payment ratio – a composite measure of patient complexity or case mix – CJR patients in elective MS-DRG 470 episodes were less complex in the intervention than the baseline, relative to control group patients (Exhibit 51). For both the CJR and control groups, a larger share of hospitals had predicted payment ratios greater than one, meaning patient complexity increased for both CJR and control group hospitals from baseline to intervention, although mandatory CJR hospitals experienced a smaller increase relative to control group hospitals. This is illustrated in the top left distribution in
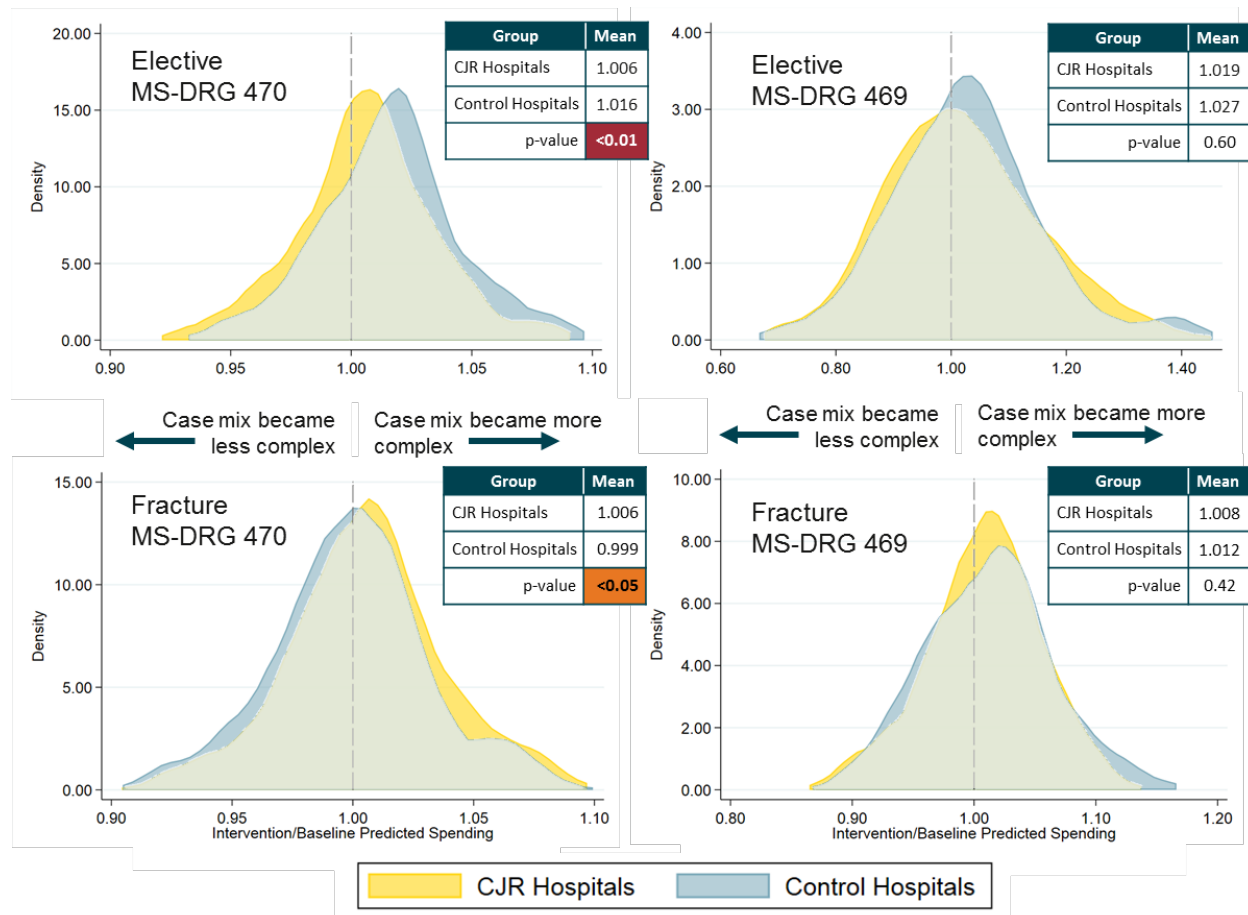
Exhibit 51; the yellow CJR distribution lies noticeably to the *left* of the blue control distribution. The mean predicted payment ratio was 1.006 for mandatory CJR hospitals, which is statistically different from the mean predicted payment ratio of 1.016 for control group hospitals (p<0.01). That is, based on changes in patient mix, we expect episode payments to increase by 0.6% for mandatory CJR hospitals and 1.6% for control group hospitals.

For elective MS-DRG 469, the distributions of predicted payment ratios for the CJR and control groups are more similar and the means for the CJR and control groups (1.019 and 1.027, respectively) are not statistically different from one another (p=0.60).

For fracture MS-DRG 470, an episode group in which hospitals have relatively limited opportunity to select patients, the distribution of predicted payment ratios for mandatory CJR hospitals lies partially to the *right* of that for control group hospitals, indicating that mandatory CJR hospitals experienced a *greater* increase in patient complexity than control group hospitals. The means of the two distributions are statistically different from one another (p<0.05) with the mean for mandatory CJR hospitals *greater* than the mean for control group hospitals (1.006 and 0.999, respectively).

For fracture MS-DRG 469, the distributions of predicted payment ratios for the CJR and the control groups are more similar and the means for CJR and for the control group (1.008 and 1.012, respectively) are not statistically different from one another (p=0.42).

Exhibit 51: Distribution of predicted payment ratios by episode group, mandatory hospitals

Source: CJR evaluation team analysis of Medicare claims and enrollment data for episodes initiated in 2012 through 2014 that ended between April 2012 and March 2015 (baseline) and episodes initiated during or after April 2016 that ended by December 2018 (intervention).

Notes: Means that are significantly different at the 99%, 95%, or 90% significance level are indicated by red, orange, or yellow shaded cells, respectively.
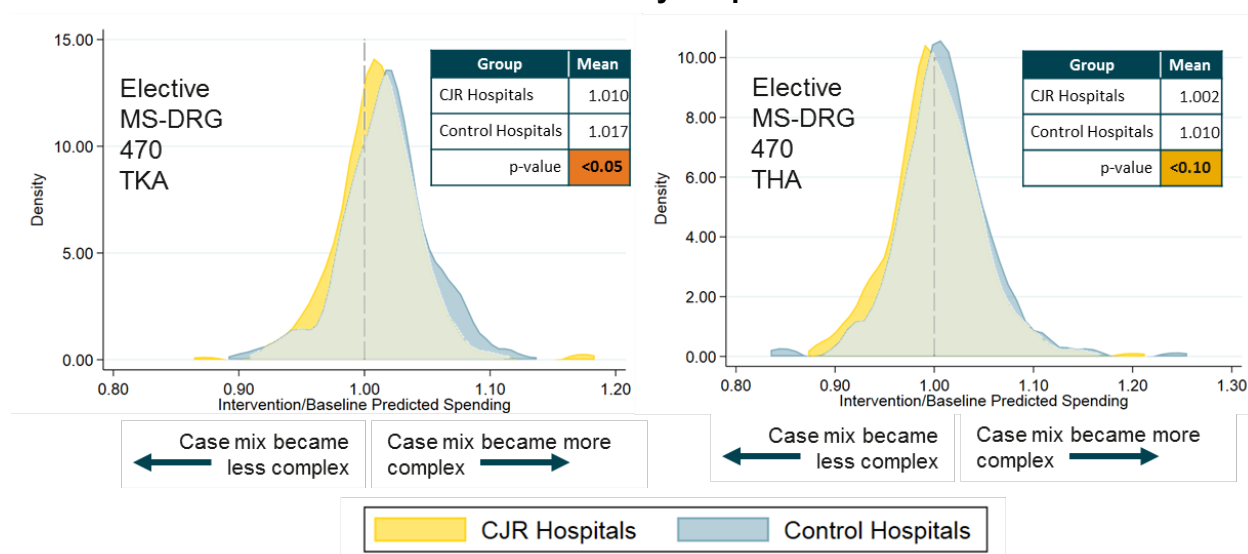
The predicted payment ratio is a hospital-specific composite measure of case mix which measures changes from the baseline to the intervention. Distribution plots were trimmed at the 5th and 95th percentile.

MS-DRG = Medicare Severity-Diagnosis Related Group.

Medicare coverage of OP TKA could explain some of the decrease in mandatory CJR hospitals' average patient complexity, relative to the control group (Section II.A.2). If CJR participant hospitals retained TKA patients of lower complexity in the inpatient setting instead of shifting them to the outpatient setting, while control group hospitals shifted lower complexity patients to the outpatient setting, the average complexity of CJR participant hospitals' inpatient population could decrease relative to the control group. Indeed, in 2018, mandatory CJR hospitals performed a lower share of TKA in the outpatient setting than control group hospitals (about 19% vs. 29%). Evidence suggests that the differential uptake of OP TKA is due to the CJR model. To better understand the link between the change in TKA coverage and patient complexity, we analyzed predicted payment ratios for elective MS-DRG 470 episodes separately for TKA and THA (a procedure not affected by the change in TKA policy).

The relative decrease in the complexity of patients in elective MS-DRG 470 episodes was evident for TKA and THA procedures. The complexity of the TKA patient population increased less for the CJR group than it did for the control group (1.010 vs. 1.017, p<0.05, Exhibit 52), resulting in a relative decrease in complexity of the CJR TKA patient population. However, this was also true for THA procedures: the complexity of the THA patient population increased less for the CJR group than it did for the control group (1.002 vs. 1.010, p<0.10), resulting in a relative decrease in the complexity of the CJR THA population. Thus it appears that mandatory CJR hospitals have decreased their elective MS-DRG 470 case mix in response to the CJR model and this occurred for both TKA and THA procedures. This suggests that mandatory CJR hospitals' differential response to the change in TKA policy is not the *sole* explanation of the changes in patient mix.

Exhibit 52: Predicted payment ratio for elective TKA and THA procedures discharged under MS-DRG 470 from mandatory hospitals

Source: CJR evaluation team analysis of Medicare claims and enrollment data for episodes initiated in 2012 through 2014 that ended between April 2012 and March 2015 (baseline) and episodes initiated during or after April 2016 that ended by December 2018 (intervention).

Notes: Means that are significantly different at the 99%, 95%, or 90% significance level are indicated by red, orange, or yellow shaded cells, respectively.

The predicted payment ratio is a hospital-specific composite measure of case mix which measures changes from the baseline to the intervention. Distribution plots were trimmed at the 5th and 95th percentile.

MS-DRG = Medicare Serverity-Diagnosis Related Group, THA = total hip arthroplasty, TKA = total knee arthroplasty.

Analyses of patient characteristics. Consistent with the composite measure analyses, the analyses of patient characteristics provide additional evidence of a reduction in the complexity of CJR patients in elective MS-DRG 470 episodes, relative to the control group.
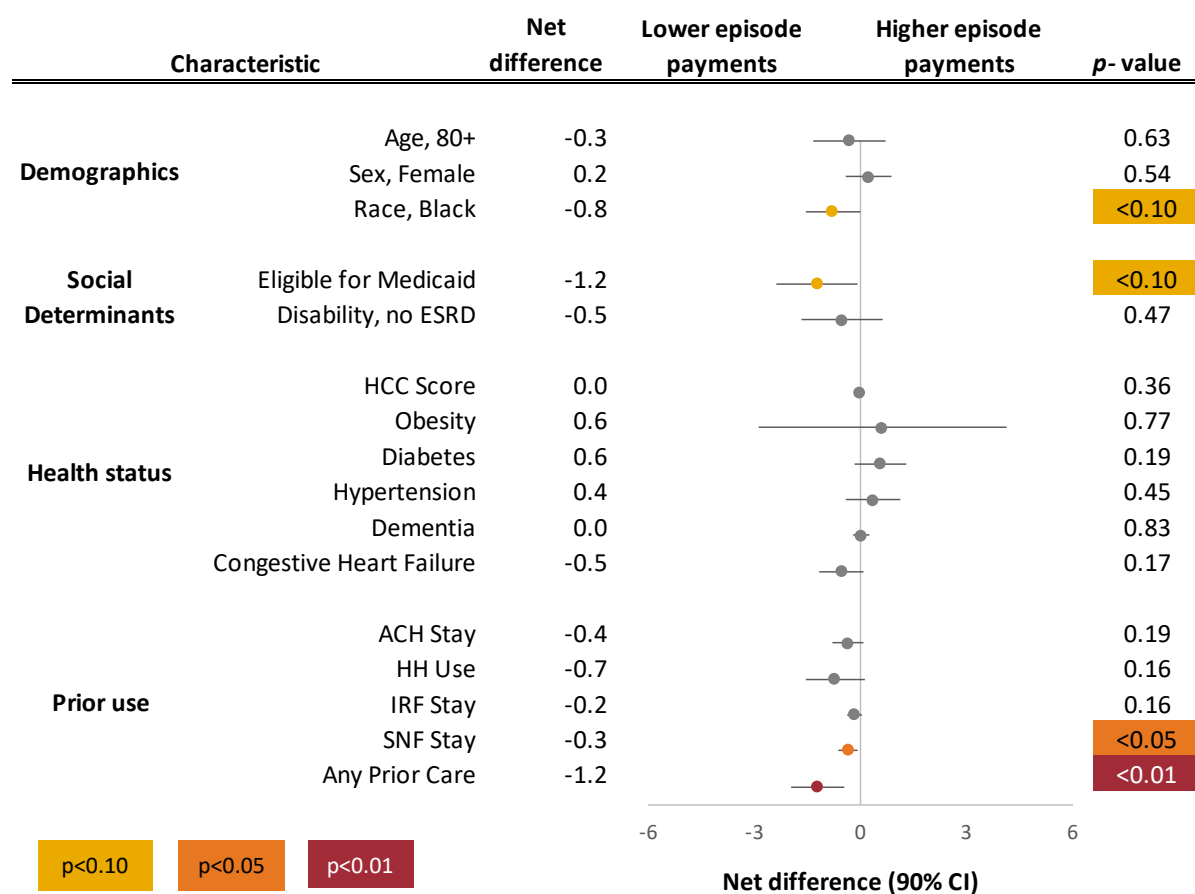
For elective LEJRs in MS-DRG 470, changes in patient characteristics consistently pointed to a reduction in the proportion of CJR patients with characteristics that are associated with higher episode payments, relative to the control group (Exhibit 53). This suggests that the CJR patient population was relatively healthier in the intervention than the baseline.

Race. There was a statistically significant 0.8 percentage point (pp) decrease in the proportion of patients who were Black (p<0.10).

Medicaid. There was a statistically significant 1.2pp decrease in the proportion of patients who were eligible for Medicaid (p<0.10).

Prior health care use. There was a statistically significant 0.3pp decrease in the proportion of patients who had a SNF stay in the six months prior to the anchor hospitalization (p<0.05) and a 1.2pp decrease in the proportion of patients who had any billed services in the six months prior to the anchor hospitalization (p<0.01).

Exhibit 53:   For elective LEJRs in MS-DRG 470 at mandatory CJR hospitals, some indications of a less complex patient population

	Characteristic	Net difference	Lower episode payments	Higher episode payments	*p*-value
Demographics	Age, 80+	-0.3			0.63
	Sex, Female	0.2			0.54
	Race, Black	-0.8			<0.10
Social Determinants	Eligible for Medicaid	-1.2			<0.10
	Disability, no ESRD	-0.5			0.47
Health status	HCC Score	0.0			0.36
	Obesity	0.6			0.77
	Diabetes	0.6			0.19
	Hypertension	0.4			0.45
	Dementia	0.0			0.83
	Congestive Heart Failure	-0.5			0.17
Prior use	ACH Stay	-0.4			0.19
	HH Use	-0.7			0.16
	IRF Stay	-0.2			0.16
	SNF Stay	-0.3			<0.05
	Any Prior Care	-1.2			<0.01

p<0.10   p<0.05   p<0.01

Net difference (90% CI)

Source: CJR evaluation team analysis of Medicare claims and enrollment data for episodes initiated in 2012 through 2014 that ended between April 2012 and March 2015 (baseline) and episodes initiated during or after April 2016 that ended by December 2018 (intervention).

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

Any prior care includes inpatient hospital, psychiatric hospital, ED visits, IRF, SNF, HH, LTCH, and hospice during the six months prior to anchor hospitalization.

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

In line with the composite measure analyses, for elective LEJRs in MS-DRG 469, the changes in patient characteristics were not consistently associated with lower or higher episode payments (Exhibit 54).
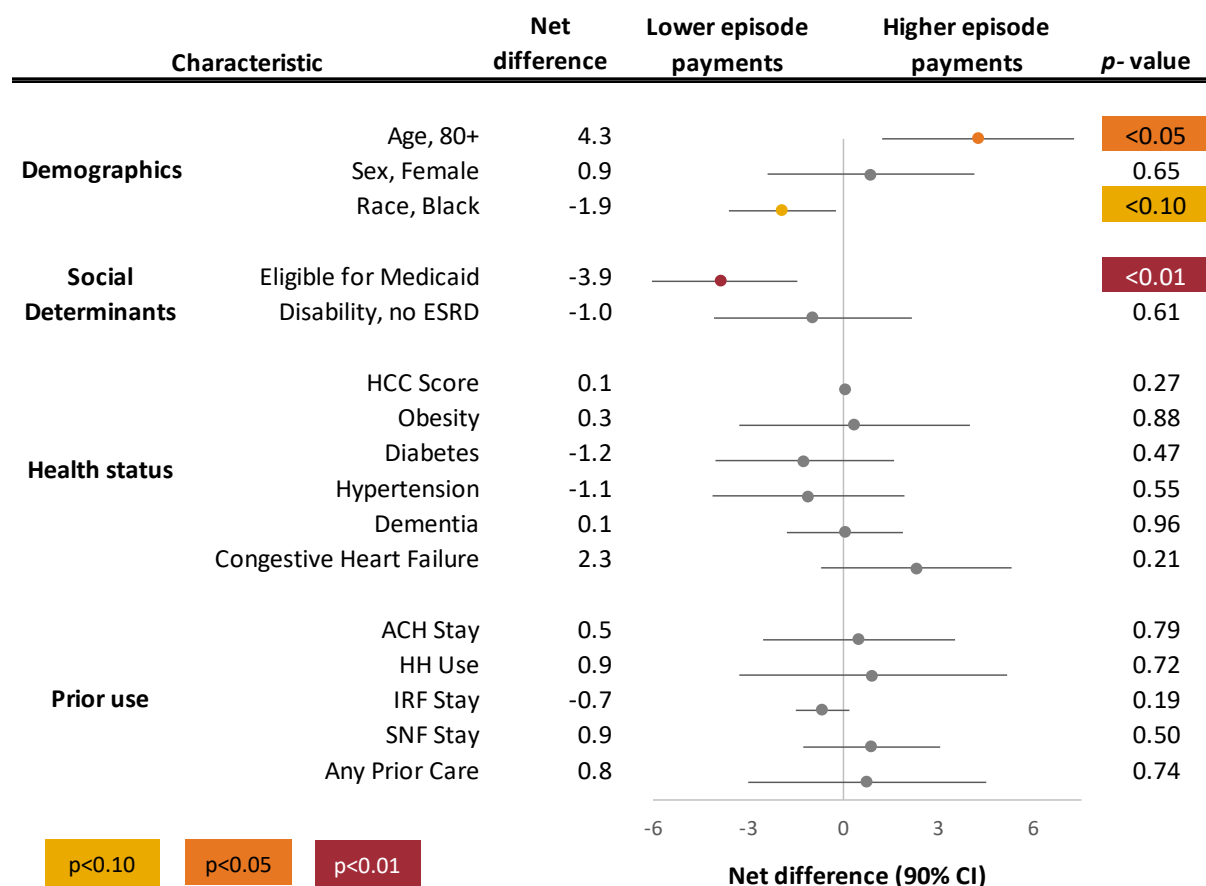
Age. There was a statistically significant 4.3pp increase in the proportion of patients who were 80 years or older (p<0.05), a characteristic associated with higher expected episode payments.

Race. There was a statistically significant 1.9pp decrease in the proportion of patients who were Black (p<0.10).

Medicaid. There was a statistically significant 3.9pp decrease in the proportion of patients who were eligible for Medicaid (p<0.01).

Changes in the remaining characteristics were not statistically significant, and the estimated changes generally had wide confidence intervals, signaling uncertainty in the estimated change. The direction of the changes was also mixed with some changes associated with lower episode payments and other changes associated with higher episode payments.

Exhibit 54: **For elective LEJRs in MS-DRG 469 at mandatory CJR hospitals, no consistent evidence to indicate a change in the complexity of the population**

Characteristic		Net difference	p-value
Demographics	Age, 80+	4.3	<0.05
	Sex, Female	0.9	0.65
	Race, Black	-1.9	<0.10
Social Determinants	Eligible for Medicaid	-3.9	<0.01
	Disability, no ESRD	-1.0	0.61
Health status	HCC Score	0.1	0.27
	Obesity	0.3	0.88
	Diabetes	-1.2	0.47
	Hypertension	-1.1	0.55
	Dementia	0.1	0.96
	Congestive Heart Failure	2.3	0.21
Prior use	ACH Stay	0.5	0.79
	HH Use	0.9	0.72
	IRF Stay	-0.7	0.19
	SNF Stay	0.9	0.50
	Any Prior Care	0.8	0.74

p<0.10   p<0.05   p<0.01

Net difference (90% CI)

Source: CJR evaluation team analysis of Medicare claims and enrollment data for episodes initiated in 2012 through 2014 that ended between April 2012 and March 2015 (baseline) and episodes initiated during or after April 2016 that ended by December 2018 (intervention).

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

c.   Conclusion

For mandatory hospitals, patient complexity increased less for CJR patients than for control group patients with elective LEJRs in MS-DRG 470, the largest patient category, thus making it easier for hospitals to receive reconciliation payments without improving the efficiency of how they provide care during the episode or improving the quality of care. This may be due, in part, to the differential shift of TKAs to the hospital outpatient setting, which may have resulted in more of the less complex TKAs remaining in the inpatient setting for mandatory CJR hospitals than the control group. The relative decline does not appear to be due to upcoding the most complex patients from MS-DRG 470 to MS-DRG 469. Nor does the relative decline appear to be due to mandatory CJR hospitals better preparing patients prior to the LEJR surgery because patient complexity differences were due to characteristics, such as race and dual eligibility status, that would not be affected by hospitals' optimizing patients before their admission.

As a result of this relative decrease in patient complexity, the CJR model may be rewarding hospitals through reconciliation payments that have not achieved its aims. Further, although we employ a rigorous risk adjustment methodology in our DiD analyses, it may not be robust enough to account for the relative decrease in CJR patient complexity for elective LEJRs in MS-DRG 470, which could bias our impact estimates and attribute greater payment reductions to the CJR model. That is, the estimated reduction in average episode payments may be overestimated because it may, in part, be due to the smaller increase in complexity of CJR patients relative to control group patients from baseline to intervention.

In addition, the relative decline in the proportion of Black beneficiaries and dually eligible beneficiaries receiving elective LEJRs raises the additional concern that existing disparities in receiving hip or knee replacements may be exacerbated under the model.[63] We will continue studying this potential impact of the model.

---

[63]   For example, see Skinner J, Zhou W, Weinstein J. The influence of income and race on total knee arthroplasty in the United States. *J Bone Joint Surg Am*. 2006;88(10):2159-2166. doi:10.2106/JBJS.E.00271 and Chen J, Rizzo JA, Parasuraman S, Gunnarsson C. Racial disparities in receiving total hip/knee replacement surgery: the effect of hospital admission sources. *J Health Care Poor Underserved*. 2013;24(1):135-151. doi:10.1353/hpu.2013.0026

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

Payments for services provided after the episode are not considered during the reconciliation process, when episode payments are assessed relative to participant hospitals' quality-adjusted target prices. While deferring services could reduce payments during the episode, this would not, however, reduce Medicare spending. Therefore, hospitals that defer services could be rewarded with reconciliation payments even though they had not achieved the goals of the CJR model.

a.   Methods

The same DiD methods described in Section II.B.1 were used for the analysis in this section. Additional details about the methodology are included in Appendix E Section III.D.1.

b.   Results

We cannot draw a conclusion about whether the CJR model affected payments for services provided during the 30 days following the episode. Although our estimate indicates that during the first three performance years the CJR model had no statistically significant impact on payments for services provided during the 30 days following the episode (DiD = -\$35, p=0.19, Appendix B, Exhibit B-1), multiple statistical tests indicate that CJR and control group hospitals did not have parallel trends for this outcome during the baseline period, so the assumptions of the statistical model are violated. Consequently, the DiD results are not reliable and we cannot draw a firm conclusion.

C. Financial Risk or Opportunity

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

After the end of each performance year, CMS reconciles each participant hospital's actual episode payments against its quality-adjusted target price. The target price is based on a blend of the hospital's average historical episode payments and its region's average historical episode payments: a hospital's target price may be lowered depending on its quality performance in the year.[64] Separate quality-adjusted target prices are set for four different episode types based on MS-DRG (469 or 470) and fracture status (elective or fracture) to account for variation in patient complexity.

If a hospital's actual episode payments are below its quality-adjusted target price and it has "acceptable" or higher quality performance, then it receives a reconciliation payment, up to a specified cap. Starting in PY2, a hospital has to repay a portion of the difference to Medicare if its episode payments are above its quality-adjusted target price. The underlying assumption of the CJR model is that the opportunity to receive reconciliation payments is an incentive for CJR

---

[64]   The "quality adjustment" occurs at reconciliation in the application of the effective discount percentage. The discount is lowered based on the hospital's quality score to incentivize quality.

participant hospitals to invest in care redesign and coordination, with the goal of increasing the efficiency and quality of care provided to patients undergoing LEJR.

In the second annual evaluation report, we identified factors correlated with earning reconciliation payments under the CJR model.[65] Hospitals that earned reconciliation payments in the first two performance years were more likely to be not-for-profit, had higher LEJR volume, had higher quality of care, and served less complex patients on average than hospitals that did not earn reconciliation payments. In this chapter, for the first three performance years of the CJR model, we identify relationships between each of these factors and the amount of reconciliation payments, while controlling for a variety of market, hospital, and patient characteristics.

---

a. Key Findings $

- In PY3, a lower proportion of CJR participant hospitals earned reconciliation payments than in previous years, which is consistent with model design changes that base a larger share of the target price on the regional average episode payments.

- After controlling for a variety of market, hospital, and patient characteristics, the average reconciliation payment per episode was higher for hospitals that had higher composite quality scores, had higher LEJR volume, were not-for-profit, and served less complex patient populations.

---

b. Methods

Reconciliation of actual episode payments against each hospital's quality-adjusted target prices is conducted each performance year of the model. Therefore, we examined reconciliation payments by hospital by performance year, allowing characteristics that may be related to reconciliation payments to vary across years. Mandatory CJR hospitals were included in this analysis if they had at least 20 episodes in the performance year. For each hospital, we calculated the average per episode amount the hospital either earned as a reconciliation payment or needed to repay to Medicare in each performance year.[66] We controlled for market characteristics (SNF beds per 10,000 population and Herfindahl-Hirschman Index), hospital characteristics (e.g., beds, ownership), and patient characteristics (e.g., age, sex, race/ethnicity, HCC score, MS-DRG, fracture status), which we averaged at the hospital level. Hospital quality performance in the CJR model, LEJR volume, and patient characteristics were performance year-specific. We categorized continuous variables based on a median split of their values to compare hospitals in the bottom half of the distribution to hospitals in the top half.

---

[65] CMS. Comprehensive Care for Joint Replacement Model. Second Annual Report. June 2019. Available at: https://innovation.cms.gov/Files/reports/cjr-secondannrpt.pdf

[66] We calculated the potential repayment amount for PY1 and used it in our analysis for consistency across performance years.

c.   Results

In the first performance year of the CJR model, 58% of mandatory CJR hospitals earned reconciliation payments (Exhibit 55). This increased to 69% in PY2 and decreased to 52% in PY3. Between the second and third performance years, the target price of 94% of mandatory CJR hospitals decreased. Because the calculation of the target price was more dependent on the regional average in PY3, average target prices for MS-DRG 470 elective episodes were $1,623 (6.3%) lower in PY3 than in PY2. As a result, hospitals needed larger decreases in episode payments to come below the target price to earn reconciliation payment or face repayment responsibility.

> The CJR model was intended to reduce variation in LEJR episode payments and reward hospitals for reducing episode payments below payments of their regional peers by using average regional target prices. To allow a transition for hospitals, in the first two performance years, two-thirds of the quality-adjusted target price was the hospital-specific average and one-third was the regional average. In PY3, one-third of the quality-adjusted target price was the hospital-specific average and two-thirds was the average regional episode amount.

Exhibit 55:   A lower proportion of mandatory CJR hospitals earned reconciliation payments in PY3 when the target price reflected a higher share of regional average payments
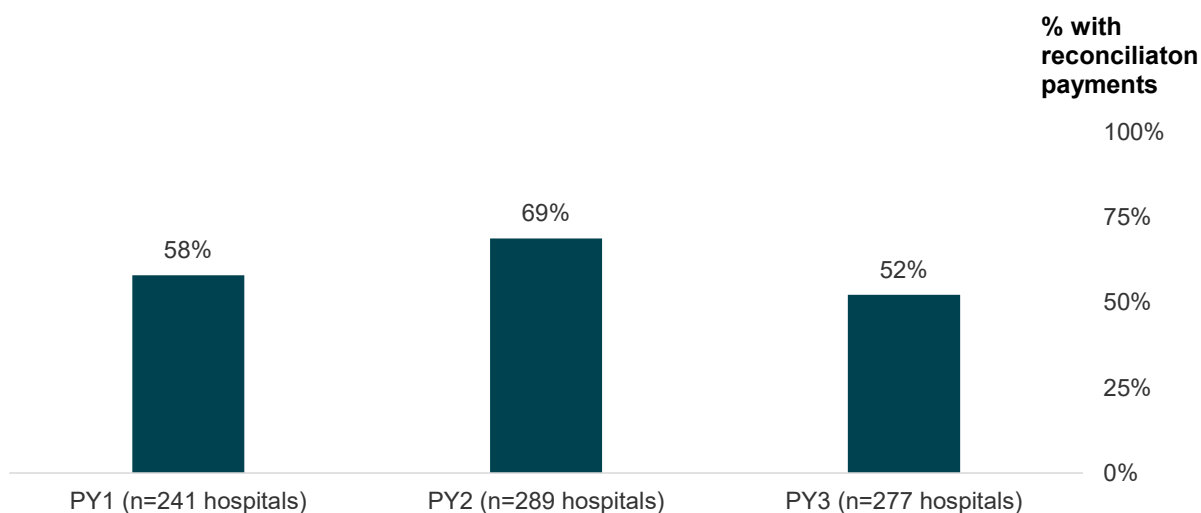
Source: CJR evaluation team analysis of CJR payment contractor data for CJR participating hospitals in PY1 (episodes starting on or after April 2016 and ending on or before December 2016), PY2 (episodes ending between January and December 2017), and PY3 (episodes ending between January and December 2018).

Notes: The sample includes mandatory CJR hospitals with 20 or more episodes in the performance year. Hospitals that did not earn reconciliation payments include hospitals with average episode payments above their quality-adjusted target prices and hospitals that were not eligible to receive reconciliation payments due to poor quality.

PY = performance year.

Factors associated with reconciliation payments

The average reconciliation payment per episode varied by performance year, quality performance, LEJR volume, hospital ownership, and hospital average patient complexity.

The average reconciliation payment for mandatory CJR hospitals was $331 higher per episode in PY2 than in PY1 (p<0.01, Exhibit 56). As reported in Section II.B.1, there was a greater reduction in average episode payments in PY2 than PY1 for the CJR group relative to the control group. The calculation of the target price changed in PY3 to be more heavily weighted to the regional average, lowering the quality-adjusted target price for most hospitals. As a result, most hospitals had to further reduce episode payments to qualify for a reconciliation payment.

Hospitals that started a performance year with average historical payments below or close to their quality-adjusted target price had higher average reconciliation payments. In a given performance year, half of mandatory CJR hospitals had average historical payments that were between 10% lower than their quality-adjusted target price and 3% higher than their quality-adjusted target price. These hospitals received $198 more per episode than hospitals with historical payments that were more than 3% greater than their quality-adjusted target price (p<0.10). This implies that some mandatory CJR hospitals may not have needed to reduce their episode payments to earn 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 with a higher quality score had a statistically significantly higher average reconciliation payment per episode. Compared to hospitals in the "below acceptable" quality category, hospitals with "good" or "excellent" quality earned higher average reconciliation payments per episode ($951 and $974 more, respectively; both p<0.01).[67]

There was a significant relationship between the hospital's LEJR volume and performance under the CJR model. Hospitals with higher LEJR volume had higher average reconciliation payments. Compared to hospitals with less than 15 episodes per quarter,[68] hospitals with 15 to 49 episodes received $499 more per episode and hospitals with 50 or more episodes received $619 more per episode in a given performance year (both p<0.01).

Further, we found that not-for-profit hospitals performed better financially under the CJR model. Not-for-profit hospitals received an average of $510 more per episode than for profit hospitals (p<0.01).

Hospitals that served less complex LEJR patients had higher average reconciliation payments. Low complexity was measured as average values below the median for patient HCC score (p<0.05), age (p<0.05), dual eligibility (p<0.05), and prior institutional PAC stays (p<0.01). In a given performance year, hospitals with lower average HCC scores (below the median of 1.59) received

---

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

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

$258 more per episode than hospitals with higher average HCC scores (p<0.05).[69] Additionally, hospitals with below the median percent of patients age 80 years or older, dually eligible for Medicaid, or with institutional PAC stays in the six months before the LEJR had higher average reconciliation payments per episode than hospitals above the median on these variables ($210, $260, and $323, respectively). (See Section II.B.6.b. for further discussion about these patient characteristics and changes in patient mix under the CJR model.)

Hospitals that served a lower percentage of low-income patients, as measured by the DSH patient percentage (below the median of 26.9%), received $243 more per episode than hospitals that served a higher percentage of low income patients (p<0.10).

Exhibit 56:   Average reconciliation payment per episode was related to performance year, quality, volume, hospital ownership, and patient complexity

Domain	Reference category	Measure	Adjusted beta coefficient	90% CI	p-value
Intercept	na	Intercept	-$2,191	-$2,550 to -$1,831	<0.01
Performance year	PY1	PY2	$331	$233 to $430	<0.01
		PY3	-$160	-$322 to $2	0.10
Difference between hospital historical average payment and PY target price	Above median (>3.0%)	Below median percent difference between hospital historical average payment and PY target price	$198	$13 to $384	<0.10
PY performance quality category	Below acceptable	Acceptable	$758	$516 to $1,000	<0.01
		Good	$951	$786 to $1,116	<0.01
		Excellent	$974	$776 to $1,172	<0.01
PY average quarterly volume	<15 episodes	15-49 episodes	$499	$322 to $675	<0.01
		50 or more episodes	$619	$407 to $830	<0.01
Ownership	For profit	Not for profit	$510	$278 to $743	<0.01
		Government	$260	-$66 to $586	0.19
Patient characteristics	Above median (>1.59)	Below median average HCC score for PY episodes	$258	$88 to $427	<0.05
	Above median (>26.6%)	Below median percent of PY episodes age 80 years or older	$210	$60 to $360	<0.05
	Above median (>14.3%)	Below median percent of PY episodes dual eligible	$260	$70 to $450	<0.05
	Above median (>5.3%)	Below median percent of PY episodes with prior institutional stay	$323	$173 to $473	<0.01

---

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

Domain	Reference category	Measure	Adjusted beta coefficient	90% CI	p-value
DSH patient percentage	Above median (>26.9%)	Below median hospital DSH patient percentage	$243	$29 to $456	<0.10

Source: Lewin's analysis of December 2016 POS, FY 2016 CMS Annual IPPS, BPCI Salesforce participation list, CJR payment contractor CJR NPRA, quality performance, Medicare claims and enrollment, and target price data for mandatory CJR participant hospital(s) in PY1 (episodes initiated during or after April 2016 that ended by December 2016), PY2 (episodes ending in 2017), and PY3 (episodes ending in 2018).

Notes: PY1 and PY2 NPRA data are final, while the PY3 NPRA data are preliminary and will be finalized spring 2020.

Generalized linear regression model, which accounts for multiple observations (PY) per hospital and clustering of hospitals at the MSA level, was used to identify factors related to average reconciliation payment per episode at the 99%, 95%, or 90% significance levels, as indicated by red, orange, or yellow shaded cells, respectively.

We restricted the sample to hospitals with 20 or more episodes in the performance year. Mandatory CJR hospitals with positive amounts per episode in a performance year received reconciliation payments under the CJR model. Hospitals with no or negative amounts per episode included hospitals with episode payments above their target price and hospitals with episode payments below their target price but with quality composite scores "below acceptable quality" making them ineligible for reconciliation payments. We calculated the potential repayment amount for PY1 because hospitals were not required to make a repayment in the first year of the CJR model. Stop gain and loss limits are applied to the overall reconciliation payment per episode.

Controlling for hospital (beds, Census region, teaching hospital, ever BPCI LEJR participation) and episode characteristics (percent female, Black, disabled, and 470 elective) (Appendix K – Factors Associated with Earning Reconciliation Payments).

BPCI = Bundled Payments for Care Improvement, CI = confidence interval, DSH = disproportionate share hospital, FY = fiscal year, HCC = hierarchical condition category, IPPS = Inpatient Prospective Payment System, LEJR = lower extremity joint replacement, MSA = metropolitan statistical area, MS-DRG = Medicare Severity-Diagnosis Related Group, NA = not applicable, NPRA = net payment reconciliation amount, POS = provider of services, PY = performance year.

d. Conclusion

While most hospitals received reconciliation payments under this mandatory model, hospitals were more financially successful in the second year of the model and those with certain characteristics earned higher payments. A higher proportion of mandatory CJR hospitals earned reconciliation payments in PY2, compared to PY1 and PY3, and the average reconciliation payment per episode was also highest in the second year. In PY3, target prices for most hospitals decreased because they were based on a larger share of regional average episode payments. As a result, a smaller proportion of hospitals earned reconciliation payments in PY3. In PY4, the target price is based completely on the regional average and the 3-year historical period for calculating the target prices includes payments under the CJR model, which will further reduce the target price for many hospitals. It will be important to continue to evaluate the target price methodology, changes in hospital patient mix, and characteristics of hospitals that are not reducing payments below their target price under this mandatory model.

Hospitals that had higher composite quality scores, had higher LEJR volume, were not-for-profit, and had less complex patient populations had higher average reconciliation payments per episode. The relationship between higher quality performance and higher reconciliation payments is consistent with model intent, to reward hospitals that focus on quality. The statistically significant relationship observed between patient complexity and reconciliation payments may indicate that

the simple risk stratification methodology used by CMS to set quality-adjusted target prices based on MS-DRG and fracture status did not adequately account for variations in patient complexity that affect episode payments within the four episode categories.

D. Resources and Market Conditions

1. How did hospital resources and market conditions influence hospitals and other providers?

As reported in the second annual evaluation report, hospital representatives said that in developing their response to the CJR model they considered their hospital's market, complete orthopedic service line, resources, and experience. It was often not possible, however, to distinguish the influence of the CJR model on changes to care pathways from broader market conditions that affected decisions about the orthopedic service line. In this third annual evaluation report, we investigate how hospital resources – such as relationships with surgeons, PAC providers, and health systems – and market conditions likely moderated hospitals' responses. We similarly found that the CJR model was only one of a variety of factors influencing interactions between CJR participant hospitals, non-CJR hospitals, orthopedic surgeons, and PAC providers.

---

a. Key Findings

Interviewees described:

■ The relationships between the hospital and orthopedic surgeons that were important in shaping hospitals' responses to the CJR model, particularly care redesign and PAC strategies.

■ Challenges associated with the low supply of quality PAC providers, including barriers to collaboration and influence over PAC provider behavior.

■ Diffusion of CJR model information, including lessons learned and best practices, through health systems to non-CJR hospitals in the same system.

---

b. Methods

We conducted site visits to hospitals and associated providers in three mandatory CJR MSAs. Our aim was to capture the effects of the CJR model on care coordination strategies, relationships with associated providers, and patient and caregiver experience from various perspectives including CJR and non-CJR hospitals, orthopedic surgical practices, post-acute care providers, and outpatient physical therapy (PT) providers throughout the MSA. We selected MSAs that varied with respect to geography, degree of provider competition (e.g. number of hospitals and LEJR volume), past BPCI LEJR participation, PAC supply and use, and outpatient TKA rate. The Provider Experiences Report provides an in-depth description of site visit hospitals' experiences under the CJR model, as well as a detailed summary of MSA-level factors that may have influenced a response to the model. We synthesized information from these site visits, as well as telephone

interviews, and clinical expert panel discussions to understand how hospital resources and market factors influenced the responses of hospitals and other providers to the CJR model. More detail on these data sources is available in Appendix E (Section I.B and Section X.C).

c. Results

Hospital relationships with orthopedic surgeons

Hospitals discussed the challenges of adapting to a risk-based model when there is a mix of employed physicians and independent physician groups who perform procedures at several different hospitals. Hospitals that did not employ surgeons often described surgeons selecting LEJR surgery sites based on personal interests and preferences, like ownership of a competing hospital or ambulatory surgical center (ASC), location, and scheduling preferences. One hospital from MSA 2 indicated that surgeons make decisions on where to perform surgery based on patient insurance, selecting one hospital for those with private insurance, another for Medicare Advantage patients, and another for Medicare fee-for-service patients who were under the CJR model. Interviewees described how this arrangement created competition between hospitals for independent surgeon's patients and often impeded the hospital's ability to influence care and drive quality outcomes. This was because hospitals were hesitant to make changes that surgeons might disagree with or see as inconvenient for fear of losing their patients to another hospital. Interviewees similarly reported that the supply of surgeons tempered the ability to redesign care in MSA 1, which interviewees described as a highly competitive market for orthopedic surgeons with the majority of surgeons employed by a single large independent physician group. One interviewee from MSA 1 stated, "That really is the challenge. They [the surgeons] are the customer, in a sense. We're trying to appease them." Interviewees from MSA 1 further described competing for surgeons with other local hospitals by offering attractive operating room schedules and catering to individual surgeon preferences for care protocols or surgical supplies to maintain LEJR volume at the hospital.

> **From the Case Studies**
>
> **CJR Hospital A** in **MSA 1** interviewees indicated that sharing cost data with orthopedic surgeons was transformative and provided the hospital with leverage to engage surgeons and change views on discharge destinations.

Telephone and site visit interviewees stressed the importance of gaining surgeon buy-in to care redesign efforts and described factors that facilitated surgeon engagement. Hospitals that employed surgeons often described success collaborating with surgeons on the design and implementation of new care redesign activities, while hospitals that did not employ orthopedic surgeons often described greater challenges. For example, one hospital indicated that the hospital's financial relationships with orthopedic surgeons contributed to their support for care redesign efforts, stating, "Surgeons at this hospital are employed by the hospital, but are also owners so they do have incentives to make sure that they lower their cost." Hospitals that did not employ orthopedic surgeons often described the importance of other "levers" that the hospital can "pull" to influence surgeon behavior, including gainsharing agreements and sharing individual surgeon performance data.

Interviewees also described that surgeon relationships with PAC providers affected hospitals' abilities to influence PAC utilization. When orthopedic surgeons either owned or had contractual relationships with PAC providers, interviewees reported it was more difficult for hospitals to control PAC use.

The level of market competition for independent surgeons and lack of "levers" limited some hospitals' response to the CJR model. For example, one interviewee explained that the hospital does not engage in patient optimization or care coordination prior to or following the hospital stay, because in the "surgeon-focused" market, hospitals compete for independent surgeons who "run the show" by making all decisions regarding LEJR patient care. The interviewee felt that employing surgeons or establishing gainsharing agreements would better position the hospital to implement care coordination strategies.

> *"You lose control, but you're dinged if they stay too long because I had a surgeon who called me…Why don't you stop this? I'm like all I can do is tell them you think the patient's ready to discharge. It's really up to that SNF director to decide that it's okay to discharge. They get very frustrated because they do lose control, but they're held accountable for it which also keeps them from sending them to SNF."*
>
> **– Hospital Interviewee**

Post-acute care market

Supply and quality of PAC providers affect how CJR participant hospitals influence changes to post-acute care pathways. Hospital interviewees identified multiple factors that limit safe patient discharges to PAC providers, including low supply, quality, bed availability, or adequate staffing levels. Many interviewees indicated that in the rural communities of their catchment areas, which had only one SNF option, they had little influence over quality of care and discharge dates, with one interviewee stating that, "Getting patients in and out of the hospital is a struggle every day."

Health system dissemination

Representatives from non-CJR hospitals indicated that they received information about the CJR model, as well as strategies and best practices in care redesign, from their health system, which included CJR participant hospitals.[70] Interviewees noted that information was shared at system-sponsored

> *"Even though we're not [in CJR], we kind of act like we are. That way, if and when we [join CJR], it would be an easy transition because we're already doing it."*
>
> **– Interviewee from non-CJR hospital**

educational forums attended by hospital and orthopedic service line leaders from both CJR and non-CJR hospitals. An interviewee from one non-CJR hospital explained that service line leaders from their health system meet every 2 to 3 months to share information, best practices, and data. The interviewee said their hospital hired a navigator after learning of the positive impacts a navigator had at one of their system's CJR hospitals. Interviewees indicated that their hospitals were receptive to information sharing and supportive of adopting strategies piloted by CJR

---

[70]   Non-CJR hospitals interviewed during site visits included hospitals outside of CJR MSAs but within the referral region of CJR participant hospitals. See Appendix E Section I.B.1 for additional information.

participant hospitals because it created efficiencies in their orthopedic care pathway while also preparing them for future episode-based payment models.

d. Conclusion

Financial arrangements between hospitals and orthopedic surgeons as well as the availability of surgeons and quality PAC providers in the market affected the level and type of control CJR participant hospitals could exert throughout the LEJR episode. When surgeons were not employed or when surgeons owned or worked at other facilities that provided LEJR, the hospital had less influence over controlling episode payments and quality. Furthermore, while the CJR model was implemented at the MSA level, health systems spanned MSAs, and interviewees reported that best practices flowed to non-CJR hospitals through system-level channels.

E. CJR Participant Hospital Responses

1. How did CJR participant hospitals respond to Medicare coverage of TKA in the hospital outpatient setting?

CMS removed TKA from Medicare's inpatient only list, effective January 2018, to allow Medicare coverage of TKAs performed in the hospital outpatient department. As a result, Medicare will pay for TKAs performed in the hospital inpatient or outpatient setting. In this section, we explore how CJR participant hospitals responded to the policy change.

---

a. Key Findings

Interviewees reported:

■ Confusion about how to interpret or respond to the rule change.

■ Changes to processes and practices for determining and documenting inpatient or outpatient status.

■ Patient care and care coordination was the same for TKA patients regardless of inpatient or outpatient status.

■ Concerns regarding the negative financial impacts of the rule change on hospital payments and performance under the CJR model.

---

b. Methods

We synthesized information from site visit interviews, telephone interviews, and the clinical review panel to understand how CJR participant hospitals responded to the removal of TKA from the inpatient only list. More detail on these methods is available in Appendix E (Section I.B and Section X.C).

c.   Results

CJR participant hospitals' responses to the policy change

To better understand variation in the share of TKA shifted to the outpatient setting, we spoke with hospital representatives about how they interpreted the policy change and whether they made any changes in response.

> **CY 2018 Final Rule:**
> "The decision regarding the most appropriate care setting for a given surgical procedure is a complex medical judgment made by the physician based on the beneficiary's individual clinical needs and preferences and on the general coverage rules requiring that any procedure be reasonable and necessary."

CJR participant hospitals' interpretation of the policy change

In the CY 2018 final rule, CMS made clear that it defers to providers to establish criteria for identifying appropriate patients for an outpatient TKA, however, lack of specific guidance regarding the application of the 2-Midnight Rule resulted in many hospitals fearing payment denials when the Recovery Audit Contractors (RACs) are allowed to review TKA patient status beginning in 2020.[71] Interviewees often indicated that the uncertainty around future payment denials influenced their response to the rule change. Some interviewees felt that the audit suspension may have hindered hospitals' responses to the rule change because it delayed feedback from CMS on the appropriateness of patient selection and documentation.

Interviewees noted confusion regarding documentation requirements and wanted clearer CMS guidance, noting that the information released by CMS was vague and did not help in interpreting the rule change. For example, interviewees said the examples provided by CMS of patients with stays lasting fewer than two midnights who might qualify as inpatients were "too open to interpretation" and insufficient in explaining CMS expectations.

> **2-Midnight Rule**
> With the rule change, inpatient TKAs may be subject to medical review by Medicare Administrative Contractors to determine if the inpatient stay was necessary. The review considers documentation in the medical record that supports either the 2-Midnight Benchmark or the Case-by-Case Exception.
>
> **2-Midnight Presumption**: Hospital claims with lengths of stay greater than 2 midnights after the formal admission are presumed to be reasonable and necessary for Medicare Part A (inpatient) payment.
>
> **2-Midnight Benchmark**: Hospital claims are generally payable under Medicare Part A if the admitting practitioner reasonably expects the beneficiary to require medically necessary hospital care spanning 2 or more midnights and this expectation is supported by the medical record documentation.
>
> **Case-by-Case Exception**: Hospital stays that are expected to span less than 2 midnights may be classified as inpatient and payable under Medicare Part A on an individualized basis if the medical record documentation supports the admitting physician's judgment that the beneficiary required hospital inpatient care despite lack of a 2-midnight expectation based on medical factors.

---

[71]   The CY 2018 final rule prohibited Recovery Audit Contractor (RAC) patient status reviews for TKA procedures performed in the hospital inpatient setting for two years (CY 2018-2019) to allow hospitals time to adopt to the rule change without fear of payment denials.

Interviewees noted that they sought additional information from industry organizations, but the information was inconsistent across sources, resulting in further confusion. For example, one interviewee described how their hospital administration's interpretation of the rule change based on CMS documents conflicted with their surgeons' interpretation, which was based on American Academy of Orthopedic Surgeons (AAOS) and American Association of Hip and Knee Surgeons (AAHKS) guidance. The interviewee said that CMS required documentation justifying inpatient procedures, while an AAHKS position statement indicated that inpatient TKA should be the norm.

Determining patient status

Interviewees most commonly identified comorbidities, age, post-acute care needs, and social supports as the key factors influencing inpatient/outpatient status determinations. Other factors reported by interviewees that influenced inpatient/outpatient decisions included patients' functional status after surgery and physical therapy needs. Several interviewees described their concern for patients who did not meet inpatient criteria based on medical need, but lacked the support or ability to go directly home post-discharge. Interviewees indicated that in these instances, hospitals would keep patients in the inpatient setting so that they qualified for a SNF stay.

> *"Outpatient patients do not qualify for the CJR 3-day SNF waiver, so the hospital feels the need to convert patients to inpatient episodes only to justify a SNF admission."*
>
> **– Hospital Interviewee**

Interviewees reported that multiple hospital staff were involved in the decision-making process including surgeons, who played a lead role in status determination, hospitalists, care coordinators, physical therapists, occupational therapists, and other medical specialists with additional knowledge about patients' comorbidities.

Roughly one third of the hospitals we spoke with assigned a default status to all TKA patients and then adjusted their inpatient or outpatient status if warranted. About half of the interviewees who discussed assigning a default status indicated that they assigned outpatient status as the default and the other half assigned inpatient status as the default. The majority of hospitals that reported assigning an outpatient default status had a short average length of stay prior to the rule change and performed a high volume of outpatient TKA, while all of the interviewees who identified a default inpatient status had a long average length of stay prior to the rule change and performed a low volume of outpatient TKA procedures.[72]

> *"The hospital does not dictate a patient's inpatient or outpatient status. Instead, we utilize a co-medical management model using hospitalists and surgeons who determine inpatient or outpatient status based on patient need."*
>
> **– Hospital Interviewee**

Interviewees from hospitals that performed fewer outpatient TKA procedures often described their patient population as medically complex and having high rates of comorbidities; most specified

---

[72] Short and long length of stay were defined as the median split in length of stay (2.4 days) for hospitals that participated in the round of telephone interviews. High and low volume were defined by a natural break in the data at 43% of Medicare TKAs performed outpatient in 2018. See Appendix E Section I.B.2 for more detail.

that their hospital had limited opportunity to perform outpatient TKA because the complex patient population was not suitable for the outpatient procedure.

Changes in documentation

Most interviewees reported that their hospital implemented changes to the documentation process in response to the rule change. Interviewees described providing more details on patient history and results of physical examinations and pre-surgical screenings in a patient's medical record. Interviewees described modifying or creating new algorithms, intake forms, or other tools to determine and then document patient status in response to the rule change. Interviewees often stated that these tools were created to establish uniform guidelines based on comorbidities and medical complexity. Hospitals often based tools on existing standards or guidelines, such as the American Society of Anesthesiologists (ASA) classification guidelines and AAHKS and AAOS guidelines, while others adopted screening tools like the Risk Assessment Prediction Tool (RAPT), and digital checklists provided by companies, such as Allscripts and InterQual, for evaluating medical necessity for inpatient surgery. Interviewees also reported making changes to their electronic medical records (EMR) systems to facilitate decision-making and documentation of patient status.

Hospitals created or modified internal audit processes to ensure adequate documentation to support inpatient status. Interviewees mentioned that case managers, utilization review staff, or anesthesiologists reviewed physicians' records to ensure documentation appropriately justified the patient's status. A handful of interviewees mentioned that additional staff, specifically, the chief medical officer, physician advisor, or medical director, reviewed the surgeons' medical notes for inpatient procedures and provided advice and guidance to surgeons on patient status determination.

Patient care and care coordination

> *"I'm doing the same operation, on the same patient, in the same setting with a significant financial penalty for that. That makes absolutely no sense to me."*
>
> **– Hospital Interviewee**

The majority of interviewees indicated that individual patient care and care coordination did not vary based on inpatient or outpatient status. Many also noted that even though the hospital provided the same care, the payment was less for outpatient procedures.[73] Interviewees commented that some outpatient TKA patients need to stay in the hospital and occupy an inpatient bed after their procedure, however, the hospital is not reimbursed at the inpatient rate. Several interviewees indicated that hospital revenues declined due to the rule change.

Interviewees described changes to care coordination efforts as a result of the rule change, including increased focus on patient screening and optimization prior to surgery (e.g., recommendations for prehabilitation). A handful of interviewees described additional hospital efforts to decrease patient

---

[73]   Based on Medicare 2018 base rate payment of $12,380 for MS-DRG 470 and $10,123 for Ambulatory Payment Category 5115.

length of stay (LOS) in response to the rule change, including changes to operating room scheduling, timing and intensity of physical therapy, and implementation of enhanced recovery after surgery (ERAS) protocols.[74]

A few interviewees described challenges with care and care coordination of patients receiving TKAs in an outpatient setting. These interviewees felt that the shorter LOS limited the "window of opportunity" for patient education and discharge planning, resulting in patients being less prepared for discharge to home. One interviewee described that "care coordinators are rushed with outpatient patients" as compared to inpatients who "spend a few days with the care coordinator discussing the discharge plan." A handful of hospitals expressed frustration that the resources they invested in care redesign and care coordination in response to the CJR model would not be recouped financially given the rule change. The interviewees often described feeling financially "punished" for successfully optimizing LEJR patients and decreasing LOS prior to the rule change.

> *"The good work that we've done has actually worked to our disadvantage because the denominator for those [CJR] indicators is now less. It's the more complex patients, and we don't get the benefit of counting those patients that we have very closely managed before surgery. When they [CMS] took it [TKA] off the inpatient-only list, I believe it was very unfair for those in the CJR program."*
>
> **– Hospital Interviewee**

Reported effect of the policy

Interviewees said the rule change will impact their CJR reconciliation payments because the quality-adjusted target price will be harder to reach as healthier patients are shifted to the outpatient setting, increasing the relative complexity of the remaining inpatients. One interviewee noted that CJR patients who were "favorable to the bundle…are no longer in [the bundle]," which the interviewee said puts the hospital "in a much compromised position." Interviewees often inquired if CMS would modify the CJR model in response to the rule change and apply an adjustment factor to the reconciliation payments to account for healthier patients no longer being included in the CJR model.

> *"When targets are established based on historical costs, and then you change the patient population by taking out the healthy ones, but the targets keep going down based on historical, it's not a real target anymore."*
>
> **– Hospital Interviewee**

---

[74] Enhanced recovery after surgery (ERAS) protocols are multimodal perioperative care pathways designed to achieve early recovery after surgical procedures by maintaining preoperative organ function and reducing the stress response following surgery.

Interviewees expressed concerns about the potential negative impact of the rule change on readmissions, "revisits," or patient satisfaction.[75] Some noted that earlier discharge from the hospital may lead to more patients returning to the hospital, which has financial implications for the hospital. One interviewee who described tracking revisits from outpatient procedures, noted that the readmission rates for inpatient procedures was similar to the revisit rate for outpatients, even though inpatient procedures were usually performed on more complex patients. One interviewee commented that the CJR quality measures used to calculate the composite quality score are based on inpatient procedures, so the denominator for the measures has changed and will include more unhealthy patients. One hospital reflected on the negative impact to its Hospital Consumer Assessment of Healthcare Providers and Systems (HCAHPs) score because the score does not capture the experience of patients with outpatient TKAs who are healthier and most likely to be highly satisfied with their care.

About a quarter of interviewees reflected on the impact of the rule change on other payers, most often stating that commercial insurers were increasingly approving TKA in the outpatient setting since CMS announced the rule change. Interviewees said that other payers were using the rule change "as justification" for limiting inpatient TKA or expanding restrictions for inpatient total hip arthroplasty (THA). Several interviewees also described concerns that other payers did not adequately consider patient comorbidities when making determinations of patient status. A few interviewees indicated that insurance companies make determinations without knowing the patient's medical history or social circumstances – an issue exacerbated by the rule change.

> *"The consequences of the rule change need to be looked at more closely. Now, private insurers, Managed Care Organizations – they don't look at the overall health of the patient or whether it's better for them to be inpatient or outpatient. They are just saying, 'Do it as an outpatient.' The rule change had a negative impact on certain groups of patients, from a healthcare perspective."*
>
> **– Hospital Interviewee**

d. Conclusion

Hospitals indicated confusion regarding interpretation of the rule change or documentation requirements for inpatient procedures and felt that more timely and clear information from CMS would be helpful. Interviewees reported a variety of strategies for responding to the rule change including establishing an outpatient default status for all TKA patients, enhancing internal documentation review and auditing processes, and modifying or creating new algorithms, intake forms, or other tools.

Interviewees reported that while the need for education for outpatients increased, patient care remained the same regardless of inpatient or outpatient status. Interviewees were also concerned

---

[75] Only patients returning to the hospital following an inpatient procedure are classified as readmissions, leading hospitals to begin to track "revisits" for patients returning to the hospital following an outpatient procedure.

about the negative impact of the rule change on patient experience, hospital payment, and reconciliation payments. We will continue to monitor how hospitals adapt to the TKA policy.

2. How did CJR participant hospitals change care coordination?

CJR participant hospitals are financially accountable for total episode payments and quality of care. The incentives of the model, therefore, encourage hospitals to coordinate care for LEJR patients across the entire episode. Participant hospitals may respond to the model by beginning care coordination activities prior to admission to the hospital and continuing after discharge from the hospital.

a. Key Findings

- Patient education, discharge planning, risk stratification, data sharing, and post-acute care coordination were key care coordination activities at CJR participant hospitals.

- CJR participant hospitals tailored care coordination approaches based on patients' needs and noted that care coordination activities generally did not vary for CJR and other LEJR patients.

- Interdisciplinary approaches to care coordination and shared responsibilities for care coordination activities were considered important across the pre-surgical, inpatient, and post-discharge periods.

b. Methods

We synthesized information from telephone interviews with care coordinators and site visits to understand how hospitals changed care coordination for LEJR patients in response to the CJR model. For the telephone interviews, we asked points of contact from a sample of hospitals to identify the individuals most knowledgeable about care coordination efforts at their respective hospitals. We conducted telephone interviews with care coordination staff (e.g., care coordinators, care navigators, social workers, nurses), although C-suite or administrative staff (e.g., Chief of Nursing, Director of Case Management) participated as well. One of the aims of the site visits was to explore hospital perspectives on the influence of the CJR model on care coordination. We discussed care coordination efforts with a variety of individuals from CJR and non-CJR hospitals, orthopedic surgical practices, post-acute care providers, and outpatient physical therapy (PT) providers. For more detailed information about these methods see Appendix E (Section I.B).

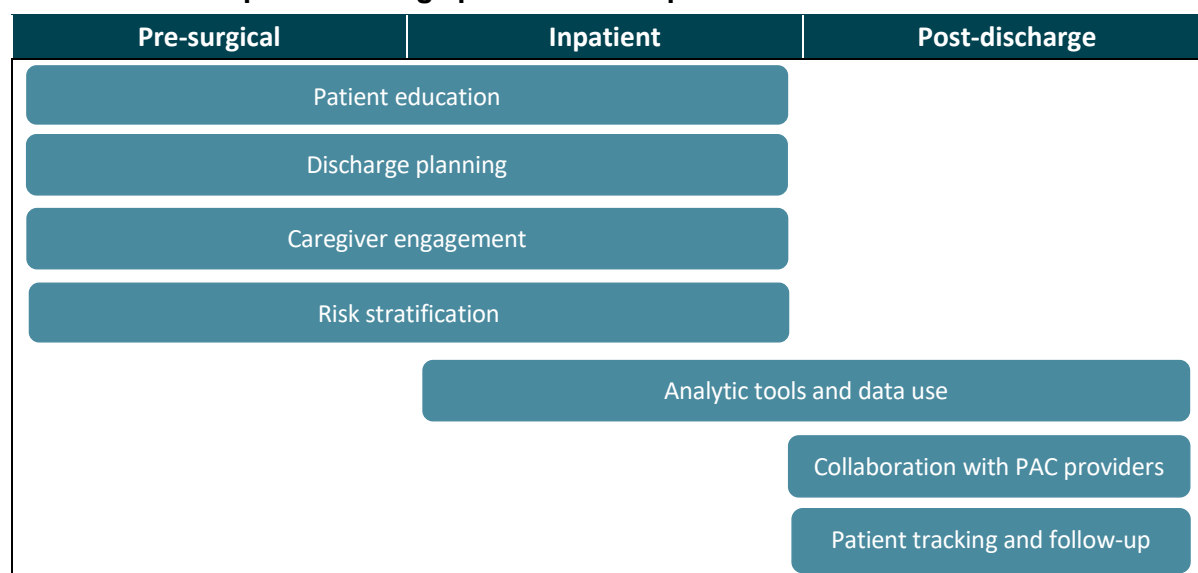### c. Results

Care coordinators from CJR participant hospitals said that participation in the CJR model impacted their preoperative classes, discharge planning, use of outcomes data, and relationships with PAC providers. Though not consistently attributed to the CJR model, interviewees used or developed tools and resources to educate patients, assess their needs, and follow them through the post-

discharge period. Interviewees described interdisciplinary approaches to care coordination, with responsibilities shared across a "care coordination team."

Care coordination across the episode

Interviewees described coordinating care for LEJR patients prior to the surgery, during the hospitalization for the LEJR, and post-discharge. Discharge planning, patient education, and risk stratification were common activities prior to and during the hospital stay. Following discharge from the hospital, interviewees reported tracking and following up with patients and collaborating with PAC providers (Exhibit 57).

Exhibit 57:   Interviewees described care coordination across the pre-surgical, inpatient, and post-discharge portion of the episode

Pre-surgical	Inpatient	Post-discharge
Patient education		
Discharge planning		
Caregiver engagement		
Risk stratification		
	Analytic tools and data use	
		Collaboration with PAC providers
		Patient tracking and follow-up

Source: CJR evaluation team analysis of telephone interview data.
Note: PAC = post-acute care.

Discharge planning and patient education

As described in previous annual reports, under the CJR model, discharge planning began earlier in the care pathway. Interviewees consistently noted that pre-surgical patient education was a key activity in their hospital's approach to earlier care coordination. Most interviewees discussed providing pre-surgical patient education through a joint camp or joint class. Classes typically covered surgery preparation and discharge planning, including expectations around discharge destination and durable medical equipment (DME) needs. Discussions about discharge destination touched on the patient's post-acute care goals and expectations, medical needs, caregiver support, environmental barriers, and unmet non-medical needs (e.g., transportation). Joint class also offered the opportunity to engage patient caregivers. Telephone and site visit interviewees noted that pre-surgical education provided in joint class could be supplemented or replaced with books, binders, or video recordings of the joint class information. These alternative options were especially useful in situations that precluded patient and caregiver joint class attendance. Additional information regarding joint class and patient or caregiver engagement is provided in Section II.E.2.

Interviewees discussed using risk stratification tools prior to the surgery to assist with identifying an appropriate discharge destination, barriers for successful transition, and risk factors for readmission. These tools were often administered during the pre-surgical clinic appointment or joint camps. If not completed prior to surgery, the assessments were administered during the hospital stay. Tools that were commonly mentioned included the Risk Assessment and Prediction Tool (RAPT tool); length of stay, acuity of admission, comorbidities, and emergency room visits tool (LACE tool); the activity measure for post-acute care (AM-PAC); and surveys or checklists created by hospitals in-house. Patients were often targeted for enhanced care coordination if they had pre-determined comorbidities, lived alone, lacked a caregiver, had unmet non-medical needs, or had limited English proficiency. Enhanced coordination activities typically included pre-surgical PT, extra education, or assessing the patient's home environment to ensure home safety and the availability of appropriate DME.

Post-discharge care coordination

Interviewees discussed post-discharge care coordination activities including patient tracking and follow-up, collaboration with PAC providers, and care coordinators' use of data from outcomes measurement and analytic activities.

The processes, timing, and duration of follow-up activities varied across hospitals. Interviewees reported that care coordinators conduct follow-up calls to monitor patients' progress after discharge from the hospital. In some hospitals staff conducted follow-up calls, while other hospitals contracted with external organizations to conduct the follow-up. Some hospitals initiated follow-up phone calls within 24 hours of discharge and others began calling patients at two weeks post-discharge. Some hospitals continued follow-up through the 90-day episode. Others described "front-loading" follow-up calls soon after discharge, which they indicated helped calm patients' nerves by addressing their concerns and demonstrating that support was available. Other hospitals did not follow-up with patients and reported placing the responsibility on patients and caregivers to reach out as needed. Follow-up data was tracked through a variety of methods, including specialized software, a homegrown spreadsheet, or the hospital's electronic system.

As discussed in previous evaluation reports, many hospitals had preferred provider lists prior to the CJR model or described updating their selection process under the model. Interviewees continued to describe collaborating or communicating with PAC providers and several noted partnerships with preferred PAC providers. Interviewees discussed ensuring SNFs understood hospital expectations and goals for patients, implementing regular meetings between hospital staff and PAC providers, or sharing claims data or electronic medical record data. Other hospitals did not expect SNFs to communicate patient-level outcomes or challenges.

Variation in care coordination activities

Interviewees commonly reported that care coordination activities did not vary between CJR and other LEJR patients. Instead, changes to care coordination activities were implemented uniformly for all LEJR patients. They did note, however, that certain activities like collection of patient-reported outcomes (PRO) data, follow up through the 90-day post-discharge period, and mandatory attendance at pre-surgical joint class were specific to CJR patients. Interviewees also discussed focusing on CJR patients during interdisciplinary meetings or when coordinating with PAC providers.

> *"If it's [pre-surgical education] such a valuable thing, why give it only to Medicare patients? We opened it up to all of our patients because every patient deserves to have that comfort level to make sure their experience is good."*
>
> **– Hospital Interviewee**

While most interviewees did not identify any differences between care coordination activities for patients with hip replacements and those with knee replacements, some described differences in patient education, with separate joint classes or education materials tailored to either hip or knee replacement. Interviewees, however, did emphasize differences in care coordination activities for patients with an elective hip replacement and those whose hip replacement was due to a fracture, noting challenges with patient and caregiver education and discharge planning for fracture patients because there was no opportunity to engage the patient prior to surgery. Some interviewees described providing alternative education or resources to fracture patients during the hospital stay. Interviewees highlighted the potential vulnerabilities of fracture patients and implications of less discharge planning, such as more intensive and longer post-acute care.

Roles and responsibilities

Interdisciplinary approaches

Most hospitals described shared responsibilities across a "care coordination team" with team members responsible for a subset of care coordination activities based on the stage in the care pathway (i.e., pre-surgical, inpatient, or post-discharge). Interviewees considered an interdisciplinary team approach a best practice, and noted that staff worked collaboratively to perform care coordination activities.

> **From the Case Studies**
>
> **CJR Hospital C** in **MSA 1** hired a total joint coordinator to drive changes that would improve performance under the CJR model. The hospital developed a "transition screening tool" that standardized the appropriate level of care following surgery, and is first used when patients meet with physical therapy representatives post-operatively.

> *"CJR has taken us from working in silos… to looking at the bigger picture of pre-op and post-discharge."*
>
> **– Hospital Interviewee**

Interviewees reported their hospitals engaged in interdisciplinary rounding, which involved collaboration among surgeons, physical therapists, and care coordinators with a goal of determining patients' post-discharge needs and appropriate discharge destination. The inclusion of patients and caregivers in interdisciplinary rounds was considered a best practice by hospitals engaged in this process.

Interviewees commonly described regular interdisciplinary meetings to discuss analysis of post-discharge data, and interviewees used analytic tools to track and share patient and hospital metrics, such as readmissions, length of stay, discharge disposition, complications, and pre-surgical class attendance. Interviewees reviewed individual patient data or aggregate patient statistics to, for example, identify ways to improve preoperative education to prevent readmissions, or identify and review successful cases with the interdisciplinary team. Consistent with prior reports, high value was placed on the receipt of CMS CJR data, and reports, dashboards, and scorecards were shared with hospital leadership and some providers (e.g., surgeons or PAC providers) to facilitate care coordination. Interviewees also used regularly scheduled meetings to disseminate findings to various stakeholder groups (e.g., quality team, interdisciplinary team, or Joint Commission).

Interviewees noted that the intensity of some care coordination activities varied depending on a patient's discharge destination. For patients discharged to a SNF, for example, there was additional coordination with the case management team at one hospital and PAC facility. Interviewees described additional coordination and communication with PAC staff as a "warm handoff."

Resources and investments

Interviewees reported that staffing and investing in care coordination was challenging. In hospitals with higher LEJR volume, interviewees mentioned challenges with high patient to staff ratios and insufficient resources to conduct all of the desired care coordination activities. For example, one interviewee noted that the ideal care coordination program would have two full-time staff members, however, this was not financially feasible. Interviewees from some hospitals said their hospitals had relatively low LEJR volume, which did not justify additional investment in care coordination. Interviewees also described how the CJR model added work to already taxed staff; one interviewee noted that a single person performed hundreds of follow-up calls with patients as an added duty due to the CJR model.

> **From the Case Studies**
>
> The not-for-profit hospital system in **MSA 1** aimed to improve care coordination by hiring a total joint coordinator and evaluating activities across the entire episode of care. The two CJR participant hospitals in this system share the total joint coordinator, a full-time position created in response to the CJR model. Both hospitals also share a data management team.

Examples of strategies hospitals employed to balance the implementation of care coordination activities with limited resources included targeting care coordination to high-risk patients, spreading care coordination expectations across hospital staff within a department to minimize burden on any one role, or co-locating services on a single hospital unit to increase communication and facilitate care coordination across providers.

Interviewees indicated that training and educating staff, for example, familiarizing them with care coordination or the CJR model, enabled them to better coordinate care for the CJR patient population. One interviewee noted that some staff completed a care coordination core competency training as part of their annual evaluation. Another noted the value of having staff that understand and can address patient concerns regarding discharge.

Care coordination metrics

Interviewees discussed measuring the effect of care coordination in terms of quality, utilization, patient satisfaction, and costs. For instance, many interviewees said they measured readmissions, infection rates, post-acute care use, discharge disposition, and length of stay. A few interviewees mentioned that they used patient surveys or patient-reported outcomes data to understand the impact of care coordination.

d. Conclusion

Interviewees from CJR participant hospitals reflected on the coordination of care for LEJR patients across the entire episode, including prior to hospital admission and after discharge, and an interdisciplinary team with duties spanning the entire care pathway was deemed a best practice. These care coordination activities, which were not consistently attributed to the CJR model, typically required significant resources, with some hospitals hiring additional staff to take on these duties while others distributed responsibilities across existing roles.

3. How did the CJR model affect patient and caregiver experience?

In response to the incentives of the CJR model to reduce episode payments, CJR participant hospitals changed PAC use by shifting LEJR patients to less intensive sites of care. These changes in the care pathway may affect patient and caregiver experiences.

---

a. Key Findings

Interviewees from hospitals and PAC providers:

- Identified challenges associated with direct discharge home and shorter SNF stays, including impacts on patient and caregiver experiences.

- Reported that setting patient and caregiver expectations during pre-surgical education helps prepare patients for earlier discharge home, which ultimately leads to better patient outcomes.

- Discussed earlier ambulation and improved pain management protocols, which support return to function and earlier discharge home.

---

b. Methods

We synthesized information from telephone and site visit interviews to describe how the CJR model impacted patient and caregiver experience. More detail on these data sources is available in Appendix E (Section I.B).

c. Results

Challenges with direct discharge home or shortened length of stay at SNFs

Telephone and site visit interviewees indicated that some patients understood the benefit of direct discharge home, while other patients, particularly older patients or patients who previously had an LEJR and expected discharge to a SNF, were more reluctant. Interviewees noted that these patients and caregivers often experience anxiety about home discharge, including addressing pain, swelling, mobilization, and activities of daily living (ADLs) without the presence of trained medical professionals typically available in SNF settings. For example, one hospital interviewee said that the biggest obstacle to discharging patients home was providing reassurance that they would be fine without a SNF admission, elaborating that patients are at risk for hospital readmission if uncomfortable being discharged home.

Interviewees stated that direct discharge home or shorter SNF stays increased reliance on caregivers. Hospital interviewees noted that it requires a significant time commitment for caregivers to provide the appropriate level of care to LEJR patients, and the caregiving responsibilities could be especially burdensome for younger caregivers who work fulltime, have children, or do not live in the same location as the patient. As one hospital interviewee shared, "It is difficult dealing with the family members. The caregiver is the one that's taking the brunt of it and they get burnt out."

Interviewees further described how caregiver support is necessary to assist patients at home with ADLs, such as toileting and bathing, noting that patients who lack a caregiver tend to have longer hospital lengths of stay and are more likely to require post-acute care than patients with available caregivers. Interviewees reported that during discharge planning, case managers verify that the patient has a caregiver who is motivated and willing to participate in the patient's care needs.

Hospital interviewees described keeping patients who would otherwise qualify for outpatient TKA in the inpatient setting longer to quality for a SNF stay if the patient lacks a caregiver.

The patient survey results indicate that returning home earlier may increase the level of caregiver support CJR patients require; however, differences between CJR and control respondents were small and there were no differences in satisfaction with overall recovery (Section II.B.5).

> **From the Case Studies**
>
> Interviewees from **CJR Hospital C** in **MSA 2** noted that they ask patients to have a caregiver present during the first 72 hours after surgery. Interviewees said that most patients have a caregiver available, however patients without a caregiver are "more likely to steer towards skilled [nursing facilities]."

A handful of outpatient physical therapist interviewees and clinical review panelists speculated that patients could incur higher out-of-pocket expenses with earlier discharges home, although this was not mentioned by hospital or post-acute care provider interviewees. Clinical review panel members noted the possibility that some patients who are discharged directly home incur costs to hire individuals to provide private care. Outpatient physical therapist interviewees also stated that copayments can be a barrier to completing physical therapy. As one interviewee stated, "I've had

patients call me and say, 'I didn't know I was going to have to pay this copay. I can't afford to keep going.' For these patients, we try to get them in for one more visit to set up a good physical therapy routine they can do at home."

SNF interviewees in certain MSAs described strong pressure from CJR participant hospitals to reduce LOS, noting that hospitals often set strict limits for members of their preferred provider

> **From the Case Studies**
>
> CJR participant hospitals in **MSA 1** described a culture before the CJR model that accepted first discharge to SNF and noted that surgeons often catered to patients' preferences and expectations of SNF use. Since the CJR model, interviewees reported having difficult conversations with patients to change expectations; however, most noted that patients ultimately agreed to a home discharge after learning of greater complication and infection rates in the SNF setting.

networks. These SNF interviewees shared that CJR participant hospitals in their MSAs often impose a seven to ten day length of stay limit for LEJR patients, but noted they discussed exceptions with hospitals as needed. One interviewee shared that this length of stay limit can be nearly "impossible" for patients with multiple comorbidities because the comorbidities "affect wound healing, their success rate, and [physical] therapy." SNF interviewees also reported some instances of patient dissatisfaction related to shortened SNF length of stay, and described an increase in the number of patients who have formally appealed their discharge date. These interviewees described increased administrative burden associated with documentation to address the discharge appeals.

Setting patient and caregiver expectations during pre-surgical education

Provider telephone interviewees shared that pre-surgical education prepares patients for discharge directly home, which ultimately leads to better patient outcomes. For example, outpatient physical therapy and home health interviewees discussed the benefit of educating patients before surgery about the time commitment and potential pain associated with physical therapy. As one interviewee reported, "Because the

> **From the Case Studies**
>
> Interviewees from **CJR Hospital C** in **MSA 2** shared that roughly 25% of CJR patients attend the preoperative class, and floor nurses "can tell a huge difference" when providing care because these patients are familiar with postoperative activities (e.g., early ambulation, home exercise programs). One respondent noted, "I wish everybody would go to preoperative class, but there could be barriers such as distance or lack of transportation."

patients are more prepared for what's going to happen, I think that contributes to a better outcome. They're more set up to know what to deal with when they go home. They do better." In telephone interviews, care coordination representatives discussed the importance of caregiver engagement during pre-surgical education, indicating that their participation reduced anxiety, especially among patients and caregivers who were nervous about being discharged home instead of to a SNF.

Some provider telephone interviewees mentioned that pre-surgical education was specific to CJR patients, while other interviewees said that the education was available to all LEJR patients, but it

was mandatory for CJR patients. In these instances, interviewees stated that CJR patients were therefore more likely to be prepared for discharge and experience better patient outcomes.

Functional status and pain

While these practices were not specific to CJR patients, site visit and telephone interviewees commonly discussed how earlier ambulation and improved pain management support CJR incentives to shift LEJR patients to less intensive sites of care and lower episode payments.

> *"About an hour after the patient gets from the operating room to the unit, they are being evaluated by physical therapy, and after that they're actually up in the room walking. Early ambulation is key."*
>
> **– Physical Therapy Interviewee**

Interviewees agreed that earlier ambulation and improved pain management support return to function and direct discharge home for LEJR patients. Interviewees also described the use of alternatives to general anesthesia as a catalyst for earlier movement right after surgery. As one physical therapist said, "It's about getting that patient up, getting them mobilized early, and getting them moving." Some interviewees identified a dedicated section on their LEJR floor where patients could practice walking, stairs, and transfer on and off a chair.

Overall, interviewees shared that changes in pain management have contributed to better outcomes for LEJR patients after discharge. Interviewees stated that surgeons have updated pain management protocols and are better educating patients about the expected level of pain after surgery. Interviewees stressed the importance of following up with patients after discharge to ensure that their pain level is under control and that they are following appropriate measures to control for pain (e.g., ice, elevation). Interviewees also shared that they call patients proactively to encourage them to take pain medication as directed.

Promising practices for supporting LEJR patients and caregivers before discharge

Although not specifically in response to the CJR model, hospital interviewees described promising practices for better understanding and addressing patient and caregiver needs. For example, one hospital interviewee described shadowing patients to better understand the LEJR experience from the patient's perspective. As a result, the hospital made care coordination changes, including streamlining patient appointments to minimize trips to the hospital. To promote early ambulation, another hospital interviewee described developing a "friendly game" where patients who received surgery on the same day compete to see who could safely walk the farthest distance. A third hospital interviewee mentioned that before LEJR patients are discharged, the hospital hosts a lunch attended by patients, caregivers, nurses, social workers, and physical therapists so that patients can socialize and learn from others who had the same procedure.

d. Conclusion

Although CJR participant hospitals and other interviewees raised a number of concerns about challenges that patients and caregivers could face with earlier discharge home, they also discussed

strategies for reducing these challenges. Interviewees described, for example, the importance of setting patient and caregiver expectations during pre-surgical education, earlier ambulation, and improved pain management protocols. Interviewees' comments were consistent with results from the patient survey, which indicate that CJR respondents required more caregiver help than control respondents, although the difference between CJR and control respondents was not large.

III. Discussion and Conclusion

A. Discussion

This third annual evaluation report on the CJR model provides additional evidence that episode-based payment and quality measurement for LEJR episodes can lower payments while maintaining quality. Mandatory CJR hospitals reduced episode payments through the third performance year while improving or preserving quality of care. After accounting for net reconciliation payments, the reductions made by the mandatory CJR hospitals resulted in net Medicare program savings, though savings decreased substantially in the third year, primarily because mandatory CJR hospitals were less likely than other hospitals to provide TKA in the hospital outpatient setting.

Hospitals responded quickly to the incentives of the model to achieve payment reductions, primarily by reducing institutional PAC use. So even though the hospitals in the CJR model might not have voluntarily chosen to participate, a range of hospitals in a variety of circumstances were able to respond. Episode payments declined in the first performance year, which was only nine months, and these declines continued through the second performance year. By the third performance year, payment reductions were moderated because mandatory CJR hospitals were less likely than comparison hospitals to shift TKAs to the lower payment hospital outpatient setting.

Even with the significant reductions in average episode payments, quality of care improved or was maintained under the CJR model. There are indications that the CJR model may be reducing unplanned readmissions. Emergency department use, the mortality rate, and manipulation under anesthesia for TKAs did not change due to the model. Further, the shifts to less intensive PAC settings did not appear to adversely affect beneficiary longer-term functional status and pain, on average. There were indications that CJR patients were less satisfied with care transitions and had greater need for caregiver help than the control patients, which could be due to their shorter SNF stays and less intensive PAC.

The CJR model is unusual among CMMI models because of its mandatory and randomized design, which minimizes the likelihood of selection bias compared with the more typical voluntary model. The change in model design at the beginning of PY3 introduced additional opportunities to test the impact of a mandatory model. The mandatory CJR hospitals in the higher payment areas, areas with greater potential to reduce spending, achieved statistically significant reductions in episode payments relative to the control group. Even with the model's more concentrated geographic focus, the variation in hospital characteristics and circumstances continues to ensure a broad test of the episode-based payment approach.

Site visits and interviews with providers involved in the entire episode of care provide insights into how hospitals shifted post-acute care from more- to less-intensive settings to reduce episode payments. Starting before the surgery was performed, hospitals indicated that they began patient and caregiver education to emphasize early ambulation and discharge home. This strategy continued in the hospital. For patients who were discharged to a SNF or home with home health, hospital representatives said they promoted early physical therapy to support shorter stays or fewer

visits. The ability to engage providers along the care continuum in efforts to reduce payments varied across participant hospitals and across markets. Hospitals that employed orthopedic surgeons had more control over care redesign and had more success in engaging surgeons than hospitals whose orthopedic surgeons worked across multiple hospitals. We found that area practice patterns, such as the continued use of general anesthesia for LEJR, could affect hospitals' ability to discharge patients home. Further, hospitals in areas with few high quality PAC providers were less likely to report success in care coordination across the episode.

The CJR model is intended to reward hospitals that reduce episode payments below their quality-adjusted target price through reconciliation payments, however, a potential unintended consequence would be if instead of improving the efficiency of care to reduce payments, hospitals treated less complex patients who needed fewer services. In fact, we found that hospitals with lower patient complexity were more likely to receive reconciliation payments. Both CJR patients and control patients with elective LEJRs without major complications or comorbidities, the largest and least complex patient category, increased in complexity during the first three years of the model, although the increase in complexity was less for CJR than control patients. As a result, the relative complexity of CJR patients decreased. Our impact estimates control for patient complexity, but unmeasured differences could remain. In that case, reductions in average episode payments may, in part, reflect changes in the mix of CJR patients rather than improved efficiency.

B. Considerations

We have employed a robust mixed methods approach that assesses the impact of the CJR model through multiple types of analyses. This approach allows results to be triangulated across data sources and methods, with shortcomings or open questions from one analysis addressed by another. Quantitative results from claims, patient assessments, and patient surveys combined with information gleaned from site visits and provider telephone interviews provide a strong evaluation of the CJR model. Consistency across findings lends strength to our conclusions, while inconsistencies raise questions for further inquiry. The CJR model's mandatory, randomized design mitigates some of the most important concerns that have hampered the evaluation of previous, voluntary episode-based payment models including selection bias and inability to generalize.[76]

While the results in this report confirm the promise of a mandatory episode-based payment model, several considerations and caveats are important to note. Our evaluation seeks to isolate the impact of the CJR model, however, interactions between the CJR model and other CMS policies and initiatives make it challenging to do so. In response to Medicare coverage of outpatient TKAs, CJR participant hospitals shifted a lower share of TKAs to the hospital outpatient setting than control group hospitals. Further, evidence indicates that the lower share is due to the CJR model. As a result, in this report, we provide two sets of estimates of the impact of the CJR model: the first is based on all LEJRs, including outpatient TKA, which are not episodes under the CJR model, and

---

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

the second is based on inpatient LEJR episodes, all of which are episodes under the CJR model but include some inpatient TKAs that likely would have been outpatient procedures absent the CJR model. Estimates based on all LEJRs (IP+OP) likely underestimate the true impact of the CJR model for two reasons. First, they include outpatient TKAs which are not episodes under the CJR model, and therefore, likely were not subject to CJR treatment. Second, average payments for OP TKAs in 2018 were higher in CJR participant hospitals than for the control group hospitals. Estimates based on only inpatient LEJR episodes likely overestimate the impact of the CJR model because they do not fully account for the differential response to the OP TKA policy between CJR and control hospitals. Taken together, the two sets of estimates provide a lower and upper bound on the impact of the CJR model after the OP TKA policy change.

Of perhaps equal importance, hospitals that participated in the former Bundled Payments for Care Improvement (BPCI) initiative contributed nearly four times the number of episodes to the CJR group than to the control group during the performance period, despite contributing nearly the same number of episodes during the baseline. The differential contribution of episodes by former BPCI hospitals during the performance period is important because, on average, former BPCI hospitals had lower episode payments than non-BPCI hospitals in the performance period, which could contribute to an overestimate of the impact of the CJR model. This finding is in line with interviews from hospitals that stated that former participation in episode-based payment models set them up to swiftly and successfully respond to the CJR model. Further testing to quantify the influence of the differential participation by former BPCI hospitals in the CJR and control groups demonstrated that the differential participation may be overestimating the reductions in average episode payments by roughly $200 per episode. The statistical significance of the impact estimate did not change under these tests, however.

Other factors may lead to an underestimate of the impact of the CJR model. The qualitative analyses indicate that diffusion of CJR care practices to non-CJR hospitals takes place through health systems, which could contaminate our control group. Any influence of the CJR model on control group hospitals would likely result in an underestimate of the CJR model. In addition, for the hospitals in all original MSAs, some hospitals may have stopped responding to the model's incentives if they knew they were not going to opt-in, which would contribute to an underestimate of the CJR model's financial impact.

The results of the patient survey by discharge setting, which found some variation in longer-term functional recovery for patients discharged to different PAC settings, may be due to increases in patient complexity within each setting as the CJR model resulted in shifts in post-acute care from more- to less- intensive settings. Even though we adjust for observable patient characteristics, decisions regarding PAC setting may be based, in part, on patient characteristics that we do not observe. If CJR patients are more complex than control patients in ways we do not observe, estimates will be biased downward (appearing to indicate worse outcomes for CJR patients). Results of the patient survey analysis by discharge setting are still useful, however, as they establish a lower-bound estimate of the impact of the CJR model within each setting.

The analysis of the site visit and telephone interview data describe common themes interviewees discussed when asked about their response to the CJR model. For site visits and interviews, we used purposive sampling to ensure that we captured a range of information on specific topics of interest for each data collection activity. While this sampling approach generates rich data about the population of interest and identifies common themes in hospitals' responses to the CJR model, it limits generalizability to the broader population.

In 2020, CMS proposed changes to the CJR model that could affect Medicare savings. Under the current CJR model, it is likely that Medicare savings will continue to decrease into the fourth and fifth performance years as more time passes under the OP TKA payment policy and as Medicare coverage of THA expands to include the hospital outpatient department and TKA is covered in ambulatory surgery centers (beginning with PY5). In part in response to these changes, CMS published a proposed rule in February 2020 that would extend the CJR model for an additional three years and expand the CJR episode definition during the extension to include OP TKAs and OP THAs.[77] The proposed rule also would make changes to the target price calculation and reconciliation process that are intended to improve the accuracy of CJR episode pricing. Beginning with the fifth performance year, the CJR model evaluation will be affected by the Coronavirus (COVID-19) pandemic. There will be fewer CJR episodes during the pandemic as CMS issued guidance stressing the need to avoid elective surgeries. In response to the COVID-19 pandemic, CMS also issued an interim final rule in April 2020 that extends the model an additional three months and, for the determination of reconciliation payments, caps episode payments at the target price for episodes impacted by the COVID-19 pandemic, effectively eliminating repayments for these episodes.[78]

Our evaluation includes numerous outcomes, which increases the risk that some of our statistically significant findings are due to chance. The strong statistical significance of many of our results and the consistent pattern of results across outcomes implies they are unlikely to be affected by this issue.

C. Conclusion

This third annual evaluation report demonstrates that the CJR model, which holds hospitals accountable for payments and quality for an episode of care that begins with LEJR surgery, remains a promising approach for reducing episode payments. Through the third year of the model, hospitals continued to respond to its financial incentives with lower episode payments than the control group. Even though our analysis focused on the mandatory CJR hospitals, the evidence indicates that hospitals in a variety of circumstances can achieve and maintain reduced episode payments. The payment reductions are due to shifts to less intensive PAC, which may have had a

---

[77] Centers for Medicare & Medicaid Services. Proposed Rule 42 CFR Part 510. Comprehensive Care for Joint Replacement Model, Three-Year Extension and Changes to Episode Definition and Pricing.

[78] Centers for Medicare & Medicaid Services. Interim Final Rule 85 FR 19230. Medicare & Medicaid Programs; Policy and Regulatory Revisions in Response to the COVID-19 Public Health Emergency.

negative impact on some patients' perception of their care, but otherwise had no adverse impact on quality of care.

Medicare program savings per episode has been substantially below the reductions in payments. While the model was intended to reward hospitals for reductions in payments and improvements in quality by allowing them to share in the payment reductions, Medicare did not achieve per episode savings equivalent to the 3% discount to the target price. In part this was because hospitals that achieved good or excellent quality scores had lower discounts applied to their target prices. However, flaws in the target price method or the risk adjustment to the prices, or both may have contributed to the diminished Medicare program savings.

The change in coverage for TKAs to allow Medicare payment in the outpatient hospital setting reinforces the finding that hospitals can respond quickly to policy changes and that financial incentives influence those responses. Because CJR participant hospitals were less likely to perform TKAs in the outpatient setting, Medicare paid more for some episodes in these hospitals than it would have absent the CJR model. The differential response to the TKA coverage change substantially reduced Medicare savings in the third year, while also reinforcing the importance of adequate risk adjustment even under a model that is designed to minimize the ability of participants to achieve reconciliation payments by changing their mix of patients.

These indications of the importance of financial incentives in health care delivery reinforce the need for thorough evaluations of models intended to shift rewards from volume to value with careful consideration of potential unintended consequences.

In future reports, we will deepen our understanding of the impact of the CJR model by incorporating subgroup analyses, refining our estimates of Medicare program savings while adjusting for other policies that affect service use and payments, and distinguishing the impact of the changes in the target price calculation. We will be able to examine additional differences in market-level effects of the model by comparing MSAs that remain mandatory with those that have a mixture of continuing and exiting hospitals. We will continue to assess unintended consequences on volume and the adoption of practice changes, and further explore the variation in patient complexity and its impact on reconciliation payments and Medicare savings. As coverage of THA expands to include the hospital outpatient department and TKA is covered in ambulatory surgery centers, we will continue to evaluate how broader Medicare policy changes affect the impact of the CJR model.