CMS Bundled Payments for Care Improvement Initiative Models 2-4: Year 2 Evaluation & Monitoring Annual Report

Prepared for:

CMS

Prepared by:

The Lewin Group

August 2016

CMS Bundled Payments for Care Improvement Initiative Models 2-4: Year 2 Evaluation & Monitoring Annual Report

August 2016

The Lewin Group

Authors:

Laura Dummit, Grecia Marrufo, Jaclyn Marshall, Ellen Tan, Aylin Bradley, Cornelia Hall, Younyoung Lee, Jon Kelly, Megan Hyland, Rebecca Cherry, Court Melin, Brandon Maughan, Ayah Fannoun, Ashley Johnson, Gina Zurbey, Susan Joy, Saran Tucker, Sebastian Negrusa, Madison Davidson, Nina Alesci, Ian Glenn, Dan Gregory, Laura Smith, David Zhang, Alex Lampert, Nick Morgan, Greyson Gordon, Jessica Steier, Dean Farley, Qian Gu, Ian Breunig, Matt Trombley, Andrea Hassol, Christine LaRocca, Lindsay Kirsch

Lewin's address:

3130 Fairview Park Dr, Suite 500, Falls Church, VA 22042

Federal Project Officer:

Table of Contents

List of Appendices

Executive Summary

The Bundled Payments for Care Improvement (BPCI) initiative is designed to test whether linking the payments for all providers involved in delivering an episode of care can reduce Medicare costs while maintaining or improving quality of care. The Centers for Medicare & Medicaid Services (CMS) launched the BPCI initiative under the authority of the Center for Medicare and Medicaid Innovation. BPCI Awardees, which can include hospitals, physician groups, post-acute care (PAC) providers and other entities, entered into agreements with CMS to be held accountable for total Medicare episode payments. Those agreements also specify Awardees' choices among four payment models, 48 clinical episodes, three episode lengths and waiver options.[1]

The BPCI initiative is designed to reward Awardees for adopting practices that reduce Medicare payments for the bundle of services in the episode relative to a target price that CMS determines based on the provider's historical payments for the same type of episode. When Awardees' episode payments are below the target price, they may receive net payment reconciliation amounts (NPRA), which they can keep or share with their partnering providers. When Awardees' episode payments are above the target price, they may have to return amounts to CMS. Thus, Awardees have strong incentives to lower episode costs.

The three BPCI Models evaluated in this report vary as to the bundle definition and payment approach.[2] The bundle is defined as the services provided during the episode that are linked for payment purposes.

- **Model 2** has the most comprehensive bundle, which includes the triggering hospital stay (i.e., the anchor hospitalization), all concurrent professional services and post-discharge services, including hospital readmissions, delivered within the chosen episode length of 30, 60, or 90 days (with certain exclusions). Individual providers are paid on a fee-for-service basis and total episode payments are reconciled retrospectively against the established target price.

- The **Model 3** bundle includes services after the anchor hospital discharge, including professional services and readmissions within the chosen episode length of 30, 60, or 90 days (with certain exclusions). The episode starts when a beneficiary is admitted to a participating skilled nursing facility (SNF), home health agency (HHA), inpatient rehabilitation facility (IRF), or long-term care hospital (LTCH) following a hospitalization for a chosen clinical episode, or when a beneficiary is admitted to a post-acute care (PAC) setting by a physician who is in a participating physician group practice (PGP). Individual providers are paid on a fee-for-service basis and total episode payments are reconciled retrospectively against the established target price.

- The **Model 4** bundle includes the anchor hospitalization, all concurrent professional services, and any readmissions and associated professional services that occur within 30 days of discharge that are not explicitly excluded from the bundle. Awardees are

---

[1] This report describes the evaluation of Models 2-4; Model 1 is evaluated separately.

[2] Within Models 2 and 3, Awardees may select one of three options for risk track. Risk track refers to the winsorization, that is, the outliers that are excluded from the reconciliation payment calculation (1st to 99th; 5th to 95th; or 5th to 75th percentile).

paid a prospectively determined amount and they, in turn, pay the providers involved in the episode.

Participation in the BPCI initiative is voluntary for providers. CMS announced on August 23, 2011 that providers and other organizations could apply to participate in Phase 1 of BPCI, during which CMS and BPCI participants could prepare for the initiative. Phase 1 participants that were approved by CMS could then enter into an agreement with CMS and begin Phase 2, which involves the assumption of financial risk by Awardees. Awardees could enter into Phase 2 as early as October 1, 2013 for any of their chosen clinical episodes. By July 2015, all participants had to transition at least one clinical episode to Phase 2 to remain in the initiative. All participants must have transitioned all of their chosen clinical episodes to Phase 2 by October 2015, when Phase 1 ended.

The roles of the providers and other organizations that participate in BPCI are distinguished by whether the entity bears financial risk, can initiate episodes, or serves as an administrator or convener. An Awardee is a provider or other type of entity that has entered into the BPCI agreement with CMS and accepts risk. An episode-initiating (EI) provider may or may not be an Awardee, but it is the provider associated with the start of the episode. Under Model 2, an EI is a hospital or a PGP; under Model 3, it is a SNF, HHA, IRF, LTCH, or PGP; under Model 4 it is a hospital. In this report, EIs and Awardees may also be referred to as participants.[3]

The Lewin Group, with our partners, Abt Associates, GDIT, and Telligen, is under contract to CMS to evaluate and monitor the impact of BPCI Models 2, 3, and 4. This is the second of five Annual Reports that synthesizes the findings from various evaluation and monitoring activities under this contract. The quantitative analyses are based on the experience of Phase 2 participants during the first year of the initiative (episodes initiated between October 2013 and September 2014). The qualitative results reflect participants' experiences through June 2015.

This observational study has limitations. BPCI is a voluntary initiative, and the providers and organizations that chose to participate differ from other providers. Although we used provider, Medicare enrollment, and claims data to select an appropriate comparison population to infer counterfactual outcomes for the BPCI population, the characteristics we selected for matching and the specificity of the data may not adequately account for all differences between participant and comparison populations.   Further, the study reflects, at most, the first 15 months of experience of the earliest BPCI participants. Because of limited sample sizes, we are not able to report on the experience of each combination of Model/participant type/ clinical episode group. Limited sample sizes, in particular, have affected our ability to understand the differential impact of BPCI across types of participants, health care delivery, and the Medicare program (research question C

---

[3] Single Awardees are individual Medicare providers that initiate episodes and assume financial risk for episodes initiated at their institution. Awardee Conveners are parent companies, health systems, or other organizations that assume financial risk for episodes initiated at their episode-initiating bundled payment provider organizations (EI-BPPOs). EI-BPPOs are Medicare practitioners that provide care to beneficiaries. Awardee Conveners may or may not be Medicare providers or initiate episodes themselves. Facilitator Conveners are entities that serve administrative and technical assistance functions on behalf of one or more Designated Awardees (synonymous with Single Awardees, but joined under a Facilitator Convener) or Designated Awardee Conveners (synonymous with Awardee Conveners, but joined under a Facilitator Convener).

below). Therefore more study is needed before generalizing these results to other providers or the full range of clinical episodes. This study also does not account for the quarterly reconciliation amounts, so additional analyses are needed to estimate total savings to the Medicare program associated with BPCI.

Our evaluation is ongoing, however, and given the recent growth in participation, which generally is not reflected in this report, our ability to detect changes in payment or quality for additional episode groups will increase.

BPCI Model	Results
Across Models 2, 3, 4	▪ From October 2013 through September 2014, the first full year of the active phase of the initiative, 94 Awardees, which include hospitals, physician groups, post-acute care (PAC) providers and other entities, entered into agreements with CMS to be held accountable for total Medicare episode payments. Across the three Models, 130 hospitals, 63 skilled nursing facilities (SNFs), 28 home health agencies (HHAs) and 4 physician group practices (PGPs) initiated almost 60,000 episodes of care under the initiative. ▪ BPCI-participating providers tended to be larger, operate in more affluent urban areas, have higher episode costs, and differed in other ways from providers that did not participate. Many indicated that commitment from their leadership and financial investment in consultants or other resources were key to implement BPCI changes. ▪ Analyses were stratified by BPCI Model, episode-initiating provider type, and clinical episode to ensure results are meaningful, but this limits the sample sizes and possibly ability to detect statistically significant effects of the initiative. ▪ Every effort was made to ensure the adequacy of comparison groups, however given the many provider characteristics that may affect outcomes and small sample sizes, all individual results should be viewed with caution and conclusions should be based on information gleaned across multiple sources. ▪ There have been modest reductions in Medicare episode payments for select clinical episode groups with isolated instances of quality declines and fewer instances of increased quality.
Model 2	▪ Model 2 was the most widely adopted model, accounting for approximately three-quarters of the episodes and half of provider participants. The majority of episode initiators were acute care hospitals, which tended to be larger, urban and likelier to have teaching programs than non-participating hospitals. ▪ The average Model 2 participant was in five clinical episodes. Almost three-quarters (74%) of Model 2 episode initiators participated in major joint replacement of the lower extremity. Congestive heart failure was chosen by 35% of episode-initiating hospitals, chronic obstructive pulmonary disease by 26%, and pneumonia by 20%. ▪ Average standardized allowed Medicare payments for the hospitalization and 90-days post-discharge were estimated to have declined $864 more for orthopedic surgery episodes initiated at BPCI-participating hospitals than episodes initiated at comparison hospitals. This was because of reduced use of institutional PAC following the hospitalization. Beneficiaries who received their care at participating hospitals indicated that they had greater improvement in two mobility measures than beneficiaries from comparison hospitals. ▪ For cardiovascular surgery episodes, institutional PAC use declined more for BPCI than comparison populations among those with any PAC. ▪ Among spinal surgery episodes, average Medicare payments increased more for the hospitalization and the 90-day post-discharge period for the BPCI than comparison population.

BPCI Model	Results
Model 3	▪ SNFs were the most dominant participants under Model 3, followed by HHAs. Only 1 each inpatient rehabilitation facility, long-term care hospital, and PGP participated. All Model 3 episode-initiating providers participated under a Convener and over two-thirds were under one of three Awardee Conveners. ▪ The average Model 3 episode initiator participated in 19 clinical episodes, the most common of which was congestive heart failure, which was selected by 95% of episode initiators. ▪ Standardized SNF payments and SNF days for SNF-initiated BPCI episodes declined relative to the comparison group across almost all clinical episode groups. However, this did not result in statistically significant declines in total episode payments. Quality generally was maintained or improved, except in three isolated instances where BPCI participant quality outcomes declined relative to the comparison group.
Model 4	▪ Model 4 was the option with the lowest number of participants and 10 out of 20 episode initiators opted out of BPCI by the end of the study period. ▪ Orthopedic surgery and cardiovascular surgery clinical episode groups accounted for 81% of all Model 4 episodes. For the orthopedic surgery clinical episode group, there were no statistically significant relative changes in Medicare standardized allowed payments, quality, or utilization. Post-bundle payments and utilization increased and certain functional outcomes declined for the cardiovascular surgery clinical episode group relative to the comparison.

A. Analytic Framework

Three major questions provide the framework for our analytic approach. Each question is addressed separately for the three BPCI Models under this evaluation. We have aggregated episodes into larger clinical episode groups when necessary to bolster sample sizes, which reduces the specificity of our findings.

> A. *What are the characteristics of the BPCI initiative and participants at baseline and how have they changed during the course of the initiative?*

BPCI is a voluntary initiative, which includes many options for participants to choose among. A complete understanding of which entities have chosen to participate, and how they have chosen to participate, helps in assessing the impact of BPCI as well as informs CMS's decisions about potential expansions or changes to the initiative.

> B. *What is the impact of the BPCI initiative on Medicare payments and the quality of care for Medicare beneficiaries?*

BPCI is one of several CMS initiatives intended to control health care costs while maintaining or improving quality of care. Answers to this question are essential for understanding whether bundled payments can help meet CMS's goal and how bundled payments may contribute to controlling costs.

> C. *What program, provider, beneficiary, and environmental factors contributed to the various results of the BPCI initiative?*

The impact of BPCI on Medicare payments and quality are likely to vary substantially because of participant choices, case mix and market characteristics. This report includes preliminary analyses to compare performance among EIs, but because of small sample sizes, there are few reported results.

B. Data and Methods

The information contained in this report was developed using a mixed methods approach designed to incorporate multiple sources of information to generate valid results. The non-experimental quantitative study design used providers in a comparison group and their episodes, during the baseline and intervention periods, to infer BPCI episode outcomes if there had been no BPCI initiative. For quantitative analyses that relied on data available only for BPCI participants, we used pre/post comparisons to estimate change attributed to BPCI. The quantitative analyses informed qualitative data collection. Qualitative data provided context and more detailed information on BPCI participant approaches and experiences, which in turn were used to help understand quantitative results and develop additional questions.

Data sources: We used the Medicare Claims and Enrollment Database from the Chronic Conditions Data Warehouse (CCW) to construct episodes of care during the baseline and BPCI intervention periods for BPCI-participating and comparison providers. We used Medicare Provider Enrollment, Chain, and Ownership System (PECOS) files to link individual physicians to PGPs. PECOS, however, has limitations, so any attributes of or results from PGP-initiated episodes should be viewed with caution. Claims were used to create Medicare standardized allowed payment, utilization, and quality outcome measures, and to identify risk factors associated with those outcomes. Patient assessment instruments used in the different PAC settings provided information on the change in functional status. We conducted surveys of beneficiaries with BPCI episodes and a matched comparison group of beneficiaries with similar episodes at non-participating providers to assess differences in patient care experiences and functional outcomes that were not available from other data sources.

We developed qualitative data to provide context on BPCI participants, their partners, and their markets. We also used qualitative data to understand care redesign approaches, implementation strategies and challenges, and factors that may explain differential outcomes. The qualitative data were drawn from a number of sources. We conducted 23 site visits with BPCI participants and two focus groups to gather information on participants, their choices, and their experiences. We conducted quarterly interviews with a sample of Awardee representatives about issues raised in other analyses as well as exit interviews with participants that left BPCI. Awardee Implementation Protocols (IPs) submitted to CMS and other information Awardees submitted quarterly to Lewin, provided information on participant plans under BPCI and progress to date. Finally, because of the dominance of major joint replacement of the lower extremity (MJRLE) episodes in the early implementation of BPCI, we conducted a Technical Expert Panel (TEP) comprising experts in orthopedic surgery and recovery to gather information on changes in patterns of care, inappropriate patterns, patient populations susceptible to poor quality of care, and outcome measures to track in the evaluation.

Population: The BPCI intervention population included all episodes initiated in BPCI-participating providers from October 2013 through September 2014 (Model 2 n=43,461; Model 3 n=12,107; and Model 4 n=3,731). For the claims-based analyses, we created a comparison group that was similar to the BPCI providers with respect to organization characteristics, market share, and volume of BPCI qualifying admissions, market concentration and size, availability of PAC providers, and case mix for each Model and EI provider type from Medicare providers that were not in Phase 1 or 2 of BPCI. We were not, however, able to match on all key provider outcomes, such as baseline mortality, which limits our ability to detect changes in these outcomes attributable to BPCI. Matching methods

identified 15 comparison providers with similar characteristics and baseline claim-based outcomes for each Model 2, 3, and 4 BPCI provider. We then drew a sample of beneficiaries from these comparison providers that matched the Medicare Severity Diagnosis Related Group (MS-DRG) distribution and dates of service in the BPCI population.

We created episodes of care by aggregating Medicare claims for beneficiaries with the MS-DRGs associated with the EIs' selected clinical episodes, across the relevant providers and in the relevant period. We used CMS's BPCI rules to construct episodes during Q4 2010 through Q3 2014.

For the survey, we chose samples of beneficiaries to maximize our ability to understand the function and experience of particular subgroups of beneficiaries treated by BPCI-participating providers. Comparison groups of beneficiaries were sampled from providers that had characteristics similar to BPCI-participating providers with the same episodes and dates of service as the BPCI beneficiary survey samples.

Outcome measures: The outcomes highlighted in this report reflect Medicare payment, utilization, provider referrals and market share, and quality, including functional status and patient experience. Payments were standardized to remove the effects of wage differences and for teaching and other policy adjustments, and trended to 2014.

Analysis: Our analyses of payment, utilization, and claim- and assessment-based quality outcomes relied on difference-in-differences (DiD) models to evaluate outcomes of beneficiaries associated with BPCI providers compared with beneficiaries who received care from comparison providers. The DiD estimate compares the change in outcomes (between the baseline and an intervention period) for beneficiaries receiving care from BPCI-participating providers relative to that same change for beneficiaries receiving care from providers in the comparison group. The baseline period was from October 2011 through September 2012, which is one year prior to the beginning of Phase 1. The intervention period was from October 2013 through September 2014. This design accounts for differences between BPCI participants and non-participants that may influence outcomes, specifically controlling for observed time-invariant provider and market factors and unobserved time-invariant factors and trends. Multivariate regression models controlled for differences in beneficiary demographics, clinical characteristics observed before the anchor hospitalization, prior health care use, and provider characteristics that might be related to the outcome.

We used a variety of empirical model specifications including ordinary least squares (OLS) and logistic regressions, and duration and two-part models, depending on the type and characteristics of the outcome measure. Estimates discussed in the results section below are statistically significant at the 5% level, unless otherwise noted.

Because of the small number of participating PGPs (three in Model 2 and one in Model 3), we did not create a matched comparison group of PGPs. Instead, we compared the risk-adjusted mean outcomes for PGP episodes to the observed mean for all episodes of the same Model, clinical episode group, and care setting.

Analyses of survey data were based on cross-section comparisons. We calculated the risk-adjusted rate of improvement or decline in functional status from before the anchor hospitalization to the time of the survey, as well as the risk-adjusted measures of patient health care experience and overall satisfaction. The impact of BPCI was estimated by the difference in

risk adjusted rates of improvement/decline between BPCI respondents and respondents in the matched comparison sample.

We investigated market, provider, and BPCI initiative factors that may influence outcomes among BPCI EIs by comparing changes in risk-adjusted outcomes from baseline to intervention among EIs in the same Model for the same clinical group. Multivariate regression methods were used to risk-adjust for differences in patient mix across BPCI providers.

We conducted exploratory analyses of BPCI effects on provider referral patterns and concentration within markets, defined as core-based statistical areas (CBSAs). For select episodes, we examined changes in the concentration of BPCI beneficiaries discharged to PAC providers in the market and the BPCI provider's share of episodes in the market.

Site visits, focus groups, and quarterly interviews with Awardees were conducted to better understand provider experiences related to the BPCI initiative. The 23 site visits that we conducted were chosen to represent the range of Model and EI choices and to provide geographic diversity. Two focus groups were used to probe specific issues: 1) engagement of PAC providers, and 2) similarities and differences of care redesign across EIs under the same Awardee Convener. Quarterly Awardee interviews were scheduled to ensure that each Awardee was contacted at least once during the year. The information from these sources was qualitatively coded for major themes using Atlas.ti software.

C. Results

Providers and other organizations have participated in the BPCI initiative across the range of the 48 clinical episode groups, although orthopedic surgery episodes have dominated. The various analyses of BPCI participants and beneficiaries who were treated by BPCI-participating providers indicates that providers have responded to the incentives of the initiative, although statistically significant differences in episode costs or quality between treatment and comparison providers have been few because of the short period under the initiative and the time needed to redesign care and the small numbers under the various clinical episodes. In addition, individual results should be viewed with caution because we are testing a large number of outcomes, so some statistical differences may be due to chance alone.

Model Description	Participation
Model 2: Anchor hospitalization plus 30,60, or 90 days of care post-discharge, maintain FFS payments	61 Awardees 110 Hospital EIs 3 PGP EIs 42,572 episodes
Model 3: Services post-discharge for 30, 60, or 90 days, maintain FFS payments	20 Awardees 63 EI SNFs, 28 EI HHAs, 1 IRF, 1 LTCH, 1 PGP 12,107 episodes
Model 4: Anchor hospitalization plus readmissions within 30 days, prospective rate	13 Awardees 20 Hospital EIs 3,731 episodes

1. Model 2 results

There were 61 Awardees with 110 hospital EIs and 3 PGP EIs in Model 2 of the BPCI initiative by the end of September 2014. Almost one-third of Model 2 EIs were participating under two large Awardee Conveners (ACs), accounting for over 50% of Model 2 episodes. During the first year, 42,572 episodes were initiated across all 48 clinical episodes.

Participant Characteristics: Episode-initiating hospitals differ from hospitals that did not participate. The majority (85%) of Model 2 episode-initiating hospitals were not-for-profit, compared with 59% of the non-participating hospitals. Nearly all (95%) BPCI hospitals were in urban locations, compared with 71% of non-participating hospitals.[4] On average, BPCI hospitals were also larger (359 beds vs. 188 beds), had greater teaching activity (a resident-to-bed ratio of 0.18 vs. 0.06), and had almost twice as many admissions for the 48 BPCI clinical episodes in 2011 (4,060 vs. 2,140) compared with non-participating hospitals. BPCI hospitals were similar to non-participating hospitals in terms of chain status (49% vs. 52%) and disproportionate share percent (30% vs. 28%). The two groups differed in total Medicare inpatient days, with a lower Medicare share for BPCI hospitals than non-participants (37% vs. 41%).

Hospital EIs tended to be located in CBSAs that had multiple competing providers, with none of them dominating the market, while markets without BPCI-participating hospitals tended to have fewer hospitals that had larger market shares. BPCI hospitals were also located in more densely populated areas with higher median incomes, compared with markets without BPCI participants. Markets with BPCI-participating hospitals tended to have more primary care physicians, specialists, and nurse practitioners for their populations than markets without BPCI participants, although BPCI markets tended to have fewer SNF beds. BPCI markets had higher Medicare Advantage penetration compared with non-BPCI markets (26.9% vs. 17.6%). The proportion of residents aged 65 years and over was similar between the two types of markets.

According to data collected from all Model 2 EIs, few had prior experience with bundled payments (16.8%). More of them had familiarity with pay for performance (67.3%) and shared savings (48.5%) models.

Model 2 Awardees reported several reasons for participating in BPCI, including to learn about payment reform, to pursue financial rewards, in response to urging of a Convener, to make quality improvements, and to align with existing expertise. Hospital EIs also indicated that they chose to participate under Model 2 instead of Model 4 because they believed that including PAC in the bundle increased their opportunities for reducing costs, and that maintaining FFS payments minimized their risk and their need for infrastructure changes.

Each of the 48 clinical episodes was chosen by some Model 2 Awardees. The average Model 2 EI participated in 5 clinical episodes. The most common episode was MJRLE, which was selected by 81 hospital EIs (74%) and accounted for 17,004 of Model 2 hospital-initiated episodes (44%). Participants indicated that they chose MJRLE because this procedure is generally elective and planned with a fairly standardized course of treatment. Congestive heart failure was selected by

---

[4] Non-participating hospitals are limited to hospitals included in Medicare's Inpatient Prospective Payment System (IPPS).

35% of EIs, chronic obstructive pulmonary disease by 26% of EIs, and pneumonia by 20% of EIs. Twenty-eight clinical episodes were selected by fewer than 10% of EIs.

Awardees took different approaches with respect to how they related to their EIs. Some provided resources and guidance to their EIs. Others assumed full responsibility for care redesign, data analysis, and patient management. Facilitator Conveners (FCs), which do not bear risk under BPCI, typically served administrative and technical assistance functions on behalf of Awardees, although some indicated that they made recommendations about episode and waiver selection based on data analysis. According to FCs that we interviewed, their participation in BPCI was primarily motivated by a desire to increase their expertise with alternative payment models, with some citing an aversion to financial risk as the reason for joining as an FC rather than as an Awardee.

Under BPCI, Awardees can request several waivers of Medicare requirements and certain requirements associated with furnishing telehealth and home visit services were waived for all participants. Based on an analysis of data that we received from Awardees, 43% of EIs had approval to provide beneficiary incentives during the third quarter of 2014, although only 6% actually exercised this option. The gainsharing waiver, which allows Awardees to share NPRA or other internal cost savings (ICS) or both with partnering providers, was requested by 80% of Awardees. We do not have data on how many actually shared any savings with providers; this data will be available in future evaluations. Although 63% of EIs had approval to use the 3-day hospital waiver, which allows Medicare coverage of a SNF stay for beneficiaries following a hospitalization of less than 3 days, only 26% of EIs used it. While Awardees supported the ability to waive the 3-day stay requirement, they were concerned about using it because of difficulties in accurately identifying beneficiaries in BPCI episodes. No Model 2 participants used the telehealth or home visit waivers during the first year of the initiative.

BPCI participants indicated that they had entered into a variety of relationships with other organizations to prepare for and participate in BPCI. Most frequently, participants indicated that they had engaged external consultants to provide data analysis or information technology. According to site visit interviews, participants tried to collaborate with area providers, particularly PAC providers, in efforts to improve care coordination and gain efficiencies across the entire episode of care. There were few specific examples of successful collaborations and participants we spoke with indicated that it was challenging to establish relationships with other providers. The Awardees indicated that they discussed quality management with the PAC providers likely to receive the Awardees' patients, even if the providers were not contractually involved in BPCI.

We discerned few distinctions across participants with regard to changes they were making to respond to BPCI. According to IPs, the majority of Model 2 EIs reported that they were implementing interventions related to patient engagement, risk management, care coordination, redesign of care pathways, and enhancements in care delivery. Several participants highlighted patient education efforts in site visit interviews, which they indicated were essential to their care redesign, and many reported that they focused on reducing PAC costs. Among the BPCI challenges participants identified were managing patient expectations related to PAC use, increasing care standardization, and accurately identifying patients who were in BPCI episodes.

Impact of BPCI: While we did not detect statistically significant changes in Medicare standardized allowed payments or quality between BPCI and comparison group episodes from

the baseline to the intervention period for most clinical episodes groups, total standardized payments declined for clinical groups that constitute most of BPCI episodes. In this section, we highlight the key findings for the three clinical episode groups where we observed a statistically significant (at the 5% level) change in total payments. These three clinical episode groups accounted for 22,761 episodes during the first year of the initiative, or 53% of all Model 2 episodes.

Orthopedic surgery: During the first year of BPCI, 82 hospitals (75% of Model 2 hospital EIs) participated in at least one orthopedic surgery episode. BPCI-participating hospitals initiated 18,936 orthopedic surgery episodes, of which approximately 90% were for major joint replacement of the lower extremity. PGP-initiated episodes were analyzed separately.

Average Medicare payments for the anchor hospitalization and the 90-day post-discharge period (PDP) were estimated to have declined $864 (3%) more for orthopedic surgery episodes initiated at BPCI hospitals than for orthopedic surgery episodes initiated at comparison hospitals. The larger reduction in Medicare payments for the BPCI population was primarily due to reduced payments for institutional PAC.

The changes in Medicare payments for orthopedic surgery episodes were consistent with differential changes in utilization during the 90-day PDP. The average SNF length of stay (LOS) among beneficiaries with any SNF use was 1.3 days shorter during the intervention period than the baseline period for beneficiaries discharged from BPCI-participating hospitals, a statistically significant decline relative to the LOS for those discharged from comparison hospitals, which remained virtually unchanged. Additionally, there was a greater decline in institutional PAC (SNF, IRF, or LTCH) use for the BPCI population than the comparison population. Among BPCI beneficiaries who received any PAC, the share that had institutional PAC use decreased from 64% to 57% between the baseline and intervention periods. This decline in institutional PAC use was 4.9 percentage points greater than the decline among beneficiaries in the comparison population (63.2% to 61.2%).

Quality of care appeared comparable between orthopedic surgery patients with BPCI episodes and those with episodes at comparison hospitals. There were no statistically significant differences between the populations in hospital readmission rates and emergency department visits within 30 or 90 days of hospital discharge and mortality within 30 days of hospital discharge. Patient assessment measures for PAC users indicate similar improvements in multiple functional outcomes for BPCI and comparison episodes, with one exception. Among beneficiaries whose first PAC setting was an HHA, the proportion of BPCI beneficiaries with improvement in upper body dressing declined from 94.3% to 93.8%, and increased from 93.9% to 95% among comparison beneficiaries.

The beneficiary survey indicated that a greater proportion of BPCI respondents with a MJRLE episode that was initiated in a Model 2–participating hospital improved in two functional measures than comparison respondents. A larger share of the BPCI respondents (65.7%) reported improved ability to walk without resting than the comparison respondents (57.5%). Similarly, 65.4% of the BPCI respondents reported improved ability to walk up and down 12 stairs, compared with 57.9% of the comparison respondents.

The change in orthopedic surgery episode payments among BPCI-participating hospitals from the baseline to the intervention period exhibited wide variation, although total episode payments

declined for 89% of the BPCI hospitals. Total per-episode payments for the anchor hospitalization and services through the 90-day PDP for orthopedic surgery episodes declined $2,137 at the average BPCI-participating hospital. The change in total payments for episode initiators from the baseline to the intervention period ranged from a decline of $7,867, to an increase of $4,163. BPCI hospitals with the greatest declines were likelier to have obtained the hospital 3-day waiver and be located in areas with fewer SNF beds per population. Given the importance of the shift away from institutional PAC to the greater decline in total episode payments for BPCI episodes relative to comparison provider episodes, these provider characteristics may be indicators of the hospitals' focus on controlling institutional PAC.

We examined separately orthopedic surgery episodes initiated by the three PGP EIs in Model 2.[5] We did not create a comparison group of PGP providers and episodes; rather we compared PGP-initiated episodes with all BPCI orthopedic surgery episodes. Approximately 4.5% of all Model 2 orthopedic surgery episodes were initiated by PGPs and these episodes were similar to all Model 2 orthopedic surgery episodes in terms of payment, utilization, and quality of care. PGP-initiated episodes had statistically significant higher total payments for HHA services ($2,523) in the intervention period than all Model 2 orthopedic surgery episodes ($1,996) because more of the PGP-initiated episodes had HHA as the first PAC setting post-hospital discharge. Hospital readmission rates and emergency department visits within 90 days of the anchor hospitalization discharge and mortality within 30 days of the anchor hospitalization discharge were not statistically different between PGP and all Model 2 orthopedic surgery episodes.

Cardiovascular surgery: During the first year of BPCI, 30 hospitals (27% of Model 2 hospital EIs) participated in at least one cardiovascular surgery episode. BPCI-participating hospitals initiated 2,859 cardiovascular surgery episodes. There was no statistically significant difference in the change in Medicare payments for the index hospitalization and the 90-day PDP between the BPCI and the comparison episodes. For a subset of episodes, however, there was a statistically significant difference between BPCI and comparison episode payments. For 30-day cardiovascular surgery episodes with PAC use, total payments for the episode was estimated to have declined $4,149 more for BPCI episodes than for episodes initiated at comparison providers.

Similar to what we saw with orthopedic surgery episodes, institutional PAC use declined more in the BPCI episodes than in the episodes initiated by comparison providers. Among beneficiaries with a cardiovascular surgery episode who received any PAC, the share discharged to institutional PAC decreased from 55.1% to 44.2% in the BPCI episodes, which was statistically different from the 47.2% to 46.2% decline in episodes of comparison providers. For home health users, there was a statistically significant increase (1.5 visits) in the number of home health visits in BPCI episodes relative to those in comparison episodes.

There were no statistically significant changes in hospital readmissions within the 30-day or 90-day post-discharge periods, or any of the assessment-based quality measures between the BPCI and comparison populations. Although emergency department use increased more for the BPCI than the comparison population during the 30-day post discharge period, there was no difference

---

[5] Please note that the data linking individual physicians with their PGP contained errors, so these results should be viewed with caution.

in the change during the 90-day period. There was, however, a statistically significant increase in mortality for beneficiaries with cardiovascular surgery episodes in BPCI-participating hospitals relative to comparison hospitals. This result was due to an increase in the mortality rate during the 30-days post-discharge (1.6% to 1.9%) for beneficiaries in BPCI episodes at the same time that there was a decline in the mortality rate for beneficiaries with episodes in comparison hospitals (2.1% to 1.4%). Because limited sample size did not allow us to match comparison and BPCI hospitals on baseline mortality rates, this outcome may be due to underlying provider differences not related to the initiative.   (More recent results that incorporate an additional nine months of data did not indicate any statistically significant change in mortality.)

Spinal surgery: During the first year of BPCI, 20 hospitals (18% of Model 2 hospital EIs) participated in at least one spinal surgery clinical episode and initiated 966 spinal surgery episodes. Average Medicare payments for the index hospitalization and the 90-day PDP were estimated to have increased $3,477 more for spinal surgery episodes initiated at BPCI hospitals than for spinal surgery episodes in the comparison group. Payments for the anchor stay increased for the BPCI episodes, but there was no statistically significant difference in the change in Part A payments by setting or Part B payments by service relative to comparison episodes, with one exception. Part B payments for imaging and lab increased $53 more in BPCI episodes than in comparison episodes.

Mortality rates within 30 days post discharge for beneficiaries with spinal surgery episodes at BPCI hospitals declined from the baseline to the intervention period. This change in mortality for BPCI episodes was statistically significantly different from the experience for comparison episodes. Again, however, this result should be viewed with caution because of differences in baseline mortality rates between comparison and BPCI-participating providers. (More recent results that incorporate an additional nine months of data did not indicate any statistically significant change in mortality.) There were no statistically significant changes in hospital readmissions, emergency department use, or any of the assessment-based quality measures.

2. Model 3 Results

There were 20 Model 3 Awardees and 94 EIs by the end of September 2014, comprised of 63 SNFs, 28 HHAs, 1 IRF, 1 LTCH, and 1 PGP. All Model 3 EIs participated under a Convener and over two-thirds of the EIs participated under one of three Awardee Conveners. Collectively, these three Awardee Conveners accounted for 74% of the Model 3 episodes.

Participant Characteristics: SNFs and HHAs that are participating in BPCI as EIs under Model 3 differ from non-participating SNFs or HHAs that discharged Medicare beneficiaries in the same clinical episodes.[6] SNF Model 3 EIs were more likely to be for-profit than non-participating SNFs (83% vs. 69%). All of the participating SNFs were in urban locations (vs. 69% of non-participating SNFs) and fewer participating SNFs were part of a chain than non-participating SNFs (17% vs. 23%). Participating SNFs were more likely to be located in a CBSA with an IRF than non-participating SNFs (56% vs. 29%). Participating SNFs were larger than non-participants (148 beds

---

[6] Due to the small number of IRFs, LTCHs, and PGPs, we do not compare them to the universe of IRFs, LTCHs, and PGPs in this report.

vs. 110 beds). Participating SNFs were similar to non-participating SNFs in market share, five-star quality rating, and average market size.

Market competition for SNFs in CBSAs with BPCI-participating SNFs was higher than in CBSAs without BPCI-participating SNFs. BPCI-participating SNFs were located in more densely populated areas, with higher average Medicare Advantage penetration and higher median household income, as well as more primary care physicians, specialists, and nurse practitioners for their populations compared with non-BPCI markets.

The majority of HHA Model 3 EIs (93%) were for-profit entities, compared with 79% of non-participating HHAs. Participating HHAs were likelier to be part of a chain than non-participants (86% vs 28%) and had more employed nurses, an indication that they were larger.

Market competition for HHAs within BPCI markets was also higher than in non-BPCI markets. Relative to other HHAs, BPCI-participating HHAs were located in more densely populated areas, with higher average Medicare Advantage penetration and higher median household income, as well as more primary care physicians, specialists, and nurse practitioners for their populations compared with non-BPCI markets.

None of the SNF or HHA EIs indicated that they had prior experience with bundled payments. Roughly a quarter of participating SNFs had experience with pay for performance, and even fewer had experience with shared savings or other payment incentives. Just over half of participating HHAs had experience with pay for performance (53.6%), with little experience with other payment incentives or shared savings.

Model 3 EIs identified several reasons for participating in BPCI, including opportunities to learn about bundled payments and anticipated payment reform, develop innovative approaches to care, and generate financial gains. Most Model 3 EIs indicated that decisions concerning episode selection were made by the organizations' administrative leadership. Organizational expertise, episode volume, and opportunities to learn about care redesign were some of the factors that influenced episode selection. During the first year of BPCI, Model 3 EIs participated in 46 out of 48 potential clinical episodes. The average Model 3 EI participated in 19 clinical episodes. The most common clinical episode was congestive heart failure (CHF), which was selected by 95% of EIs. Some participants, though, noted challenges with the CHF episode. In interviews with participants that dropped the CHF episode, they explained that it was particularly difficult to identify BPCI episodes, manage care, and prevent readmissions for CHF patients. Chronic obstructive pulmonary disease and simple pneumonia were the next most common clinical episodes, both chosen by 74% of Model 3EIs during the first year of the initiative.

Model 3 EIs described forming or augmenting existing relationships with other PAC providers and hospitals, and engaging third-party administrators and data management contractors. These relationships were intended to help the EIs identify and track patients, improve care coordination, ensure downstream quality, and analyze performance in the initiative. While Model 3 EIs identified numerous benefits to forming strong relationships, they also noted several challenges, such as difficulties forming relationships with hospitals and physicians affiliated with different provider systems.

Many Model 3 participants indicated that they have implemented care redesign activities related to risk assessment, care coordination, and patient engagement. During site visit interviews, participants spoke about some of the barriers to implementing care redesign, such as difficulty obtaining information from hospitals to determine if the beneficiary was in a BPCI episode. The care redesign strategies for Model 3 participants focused on reducing costs by reducing PAC utilization and hospital readmissions.

Waiver use was limited among Model 3 EIs. During the third quarter of 2014, 38.3% of EIs had approval to distribute beneficiary incentives, but only half actually did so. The majority (83%) of EIs had approval to participate in gainsharing. Data on the use of gainsharing are not yet available. No participants used the telehealth or home visit waivers, which were granted to all participants.

Impact of BPCI: For most clinical episode groups, there were no statistically significant differences between BPCI and comparison Model 3 episodes from the baseline to the intervention period in total Medicare standardized allowed payments during the qualifying inpatient stay and 90-day PDP or quality measures. There were, however, statistically significant declines in the SNF payments. There were no statistically significant differences in payment for HHA-initiated episodes, with the exception of a statistically significant decrease in the total amount of payments included in the pre-bundle period for non-surgical episodes.

Standardized payments for SNF services declined in Model 3 SNF-initiated episodes, relative to the comparison group across all clinical episode groups, except non-surgical respiratory. Changes in utilization were consistent with the decline in SNF payments. For all clinical episode groups, except non-surgical respiratory, the greater decline in the number of SNF days for episodes initiated by BPCI-participating SNF EIs was statistically significant relative to SNF days for episodes initiated in non-participating SNFs.

For Model 3 SNF EIs, quality outcomes were similar to those in the comparison group, with a few exceptions. For non-surgical cardiovascular clinical episodes, there was a statistically significant increase in the unplanned readmission rate during the first 30 days of the episode, equal to 7.0 percentage points, relative to the comparison group. Also, according to patient assessment data, there was a statistically significant decline of 13.9 percentage points in the share of beneficiaries with improvement in self-care function among those with orthopedic surgery episodes, relative to beneficiaries in comparison episodes.

3. Model 4 Results

There were 13 active Model 4 Awardees in the third quarter of 2014, with 20 EI hospitals. Ten of these EIs, however, terminated their participation in BPCI during the first year of the initiative. We have included episodes initiated at the terminated EIs in our analysis up until the date at which they terminated their participation. Limited participation in Model 4 has limited our analyses and findings.

Participant Characteristics Compared with non-participating hospitals, Model 4 EIs were more likely to be non-profit (65% vs. 59%) and located in an urban area (100% vs. 71%). Half of Model 4 EIs were part of a chain, which was similar to non-participating hospitals (52%). Participating hospitals were larger (427 beds vs. 188 beds) and had more intensive teaching programs, as indicated by a higher average resident-to-bed ratio (0.14 vs. 0.06) than non-participating hospitals.

In addition, EIs had over twice as many BPCI episode admissions in 2011 than non-participating hospitals (4,516 vs. 2,140). EI hospitals averaged a lower share of Medicare days than non-participating hospitals (29% vs. 41%) and had similar Medicare disproportionate share percentages (29% vs. 28%).

Few EIs reported that they had prior experience with bundled payments (14.3%). The majority of participants were familiar with pay for performance (85.7%). Some Model 4 EIs had experience with shared savings (28.6%) or other payment incentives (28.6%).

For Model 4 EIs, the decision to participate in BPCI was largely driven by the hospitals' administrative leadership, with support from physician leaders. Model 4 participants indicated that they joined BPCI because they wanted to learn about bundled payments, as they anticipated it would be a component of future payment reform. In addition, they had identified opportunities to lower costs. We interviewed the Model 4 EIs that left BPCI. They indicated that they had faced significant challenges using the monthly data files that they received from CMS, which, in turn, delayed their payments to physicians involved in their BPCI episodes. They said this was the key reason they left the initiative.

During the first year of the initiative, Model 4 EIs participated in 17 of the 48 clinical episodes, the most common being MJRLE, which was selected by approximately 70% of EIs. Coronary artery bypass graft and double joint replacement of the lower extremity were the next most common clinical episodes, which were each chosen by 45% of EIs during the first year of the initiative.

A large majority (93%) of EIs did have approval to participate in gainsharing, but information about the use of the gainsharing waiver is not yet available.

Model 4 participants indicated that they have developed relationships with PAC providers to improve communication and care coordination. Several participants noted, though, that it is challenging for the hospital to partner with every PAC provider to which patients may be discharged. Participants indicated that they cannot track some patients after discharge, and that those patients are at an increased risk for readmission. Even when they are able to track patients, participants report that the data they rely on about readmissions is often delayed or incomplete.

In addition to efforts to improve coordination, many participants described implementing care redesign, care standardization and patient education initiatives. Some participants noted, however, that they have encountered resistance to efforts to standardize care from physicians who are reluctant to adopt new care protocols. All Model 4 participants indicated that reducing readmissions was their primary strategy to reduce costs, though they described a variety of approaches to do so.

Impact: For Model 4, we calculated the impact of BPCI on payment and quality for orthopedic surgery and cardiovascular surgery clinical episode groups. These two clinical episode groups comprised 3,021 episodes during the first year of the initiative (81% of all Model 4 episodes).

There were no statistically significant differences in the change in Medicare standardized allowed payments between BPCI and comparison group episodes for the anchor hospitalization plus the 90 days post-discharge from the baseline to the intervention period for orthopedic or cardiovascular surgery clinical episodes. Although there was a statistically significant increase in

payments during the first 30 days post-hospital discharge in the cardiovascular surgery clinical episode group, the change in mortality, emergency department visits, and unplanned readmissions between baseline and intervention periods was not statistically different between BPCI and comparison orthopedic surgery or cardiovascular surgery clinical episode groups.

There were also no statistically significant changes from baseline to intervention in assessment-based quality measures for BPCI patients with orthopedic surgery episodes compared with the comparison group. Among cardiovascular surgery patients who received their first PAC treatment at a HHA, a smaller percentage of BPCI patients demonstrated improvement in bathing, ambulation, and bed transferring relative to the comparison group. These results were statistically significant at 0.05. While the share of comparison patients that demonstrated improvement in these areas increased from baseline to intervention, the proportion of BPCI patients demonstrating improvement declined.

D.　Discussion

The BPCI initiative was designed so that participants had multiple implementation options, including the fundamental choice of whether to participate or not. Descriptive statistics reveal that providers that chose to participate in Phase 2 during the first year differ from those that delayed or chose not to participate. The hospitals, SNFs, and HHAs that participated in BPCI as EIs were more likely to be larger, urban facilities, generally located in areas with higher income populations than other providers of the same type. Taken together with information from participant interviews about preparations for entering the initiative and participants' on-going change efforts, these characteristics indicate that participants may have more financial or leadership resources to devote to responding to changing payment incentives.

During preparations for BPCI prior to Phase 2, many participants told us that consultants advised them on choices about how to participate in BPCI, including gainsharing methods, and provided on-going assistance with monitoring performance and analyzing CMS-provided data. The availability of financial resources to obtain this outside help may be an important distinguishing factor between those that participated and those that did not. In the future we will explore whether Awardees or Conveners provide the resources EIs require to participate successfully. Experience to date indicates that the Convener approaches are quite prevalent in the participation of PAC providers under Model 3.

Beyond the decision on whether or not to participate, participants must select their BPCI Model, episodes and episode lengths. Hospital EIs indicated that these decisions were based on opportunities to reduce spending and perceived risk. Often hospitals told us that they chose Model 2 because they believed they could reduce PAC use; those that entered under Model 4 indicated a lack of ability to influence PAC. As to risk, many hospital EIs we spoke with indicated that they considered the financial risk of their Model choice. Given the much larger participation in Model 2, relative to Model 4, Model 2 must have been considered financially favorable or less risky or both. The large proportion of participants in Model 4 cited information and payment issues that interfered with their ability to reimburse physicians as the main reason they withdrew from BPCI. For hospitals, episode choice was often determined by whether they had sufficient volume of cases within the episode type and the potential for savings. Many of them mentioned a physician champion in a particular specialty that affected their choice. PAC EIs under Model 3 generally chose more episode groups, possibly because of their overall smaller size. Since most of

them participated under an Awardee Convener, their choice of episodes may have been influenced by the Awardee Convener.

For most clinical episodes, there were no statistically significant differences in the change in Medicare standardized allowed payments between BPCI participants and comparison providers, although many of the participants we interviewed indicated that they had implemented efforts intended to reduce total episode costs. The lack of widespread impact of BPCI on Medicare payments across clinical episodes thus far may be due to several factors. The quantitative results are based on less than one full year of BPCI experience for the majority of participants. Although some providers have been preparing for the initiative well before entering into Phase 2, this still may not be enough time to see the impact of these changes on episode payments. There is limited sample size for most clinical episodes, which limits our ability to detect changes in payments. Further, to increase our ability to detect changes, we conducted many of our analyses on aggregations of clinical episodes, which may mask payment changes for particular types of episodes.

For Model 2 orthopedic and cardiovascular surgery episodes participants' efforts to reduce episode spending are achieving expected results. For these episodes, which account for a large share of Model 2 episodes, we saw a statistically significant shift from more expensive institutional PAC to less expensive home health care among beneficiaries discharged to any PAC setting. This shift was the major contributor to the larger relative decline in total payments during the anchor stay and the 90-day PDP for orthopedic surgery episodes. There was not, however, a statistically significant decline in payments during the anchor stay and the 90-day PDP for cardiovascular surgery episodes. Many of the Awardees and EIs we spoke with indicated that they focused their efforts on reducing PAC costs, which often involved patient education initiatives to prepare beneficiaries for a discharge home after their surgery.

The reduction in payments for orthopedic surgery episodes under Model 2 was achieved by changing the type and use of PAC. Affecting this critical decision was probably made easier by the fact that orthopedic surgery typically is elective and scheduled, allowing clinicians involved in the episode to educate the beneficiary about the best site of care. Beyond this decision, there was little evidence that EIs are affecting care after the hospital discharge. Some hospital representatives told us that they had bolstered discharge planning and case management to include periodic telephone calls with BPCI patients after their hospital discharge and to help with making follow up appointments. However, few indicated that they actively managed the care for their patients after the PAC decision.

Medicare payments for spinal surgery episodes initiated under Model 2 went up relative to payments for episodes at comparison hospitals. At the same time, there was a statistically significant decline in mortality for BPCI episodes relative to comparison episodes (although this mortality result did not persist with an additional nine months of data). These results raise questions about why payments would go up for providers that have incentives to lower total spending. These results warrant further investigation.

There was a statistically significant reduction in SNF length of stay for some Model 3 SNF-initiated episodes, although this did not reduce total episode payments. This type of change is directly under the SNF EI's control and would not necessarily require the EI to coordinate or manage care across the episode. Reducing SNF length of stay directly reduces SNF Medicare revenues, which could make this a difficult trade-off for SNF EIs. They would need to calculate

their potential financial rewards under BPCI for reducing length of stay in comparison to foregone daily Medicare payments.

The claims-based measures of quality provide no indication of systematic effects of BPCI across most clinical episode groups. For Model 2 cardiovascular surgery and spinal surgery episodes, there were statistically significant changes in mortality relative to the comparison group, although because of baseline differences in mortality rates between BPCI and comparison hospitals, these results may be more reflective of different trends between BPCI and comparison hospitals rather than differences due to BPCI. (More recent analyses indicate that there were no differences in relative mortality rates between BPCI and comparison hospitals for these two clinical groups.) Refinements to the comparison group selection methods, which will be possible as samples increase, will allow further investigation of relative changes in mortality. Further, we will examine any corroborating information that any change is related to BPCI. Unplanned readmission rates generally have been declining, although we have noted instances in which the decline was slower for Model 3 participants than their comparison counterparts. We are not able to evaluate quality or beneficiary functioning across PAC settings. Within the separate PAC settings, there were few instances in which there were statistically significant differences in the change in functioning between BPCI and comparison patients, although when there was a statistically significant difference, the BPCI patients exhibited less improvement in functioning. Particularly with the shift of beneficiaries from institutional PAC to home health care, we will continue to examine any changes in functioning in the context of likely changes in patient complexity across PAC settings. Shifts away from institutional PAC will tend to increase patient complexity in both home and institutional PAC settings.

The beneficiary survey results indicate that BPCI has not had a detrimental effect on beneficiary satisfaction with their episode of care. Even for beneficiaries with MJRLE episodes, which are within the group of orthopedic surgery episodes that exhibited relative payment declines, satisfaction with care was the same between those with BPCI episodes and the comparison providers. Further, self-reported functioning improved more for beneficiaries with BPCI MJRLE episodes in two key mobility measures. This is consistent with what many clinicians told us, that recovery from orthopedic surgery is better achieved in the beneficiary's home.

A major limitation of this evaluation has to do with small sample sizes and short exposure to the BPCI incentives for many participants. Over the years, this limitation will diminish. In addition, this is an observational study. As such, it relies on the differential change in claims-based measures between BPCI participants and a comparison group to infer the impact of BPCI. The strength of these results, therefore, is dependent on the adequacy of the comparison group in representing what would have happened absent BPCI. We continue to refine our methods for choosing a comparison group, which will strengthen our conclusions. In addition, with the large number of outcomes stratified by Model, participant type, and clinical episode, it is likely that some results will be significant by chance alone. We will continue to examine outcomes over time and use multiple methods and data sources to address research questions to be able to distinguish chance from actual results and further increase our confidence in conclusions about the initiative.

Over the next year, we will be able to examine outcomes for more disaggregated clinical groups, which will improve our ability to provide meaningful results. We will also expand our analyses of the key characteristics of participants that contribute to their ability to achieve savings and

maintain quality under BPCI. We will also be able to examine the impact of BPCI on beneficiary subpopulations.

E.   Conclusions

This second Annual Report provides a summative and formative evaluation of the BPCI initiative based on quantitative analyses of Phase 2 participants that joined the initiative during the first year and qualitative analyses of participants that joined during the first seven quarters. Most results are based on the experience of 94 Awardees across three Models, with 227 episode initiators that were responsible for 58,410 episodes of care during the first year of the initiative. We remain limited in our ability to estimate the impact of the initiative under most Model and episode combinations because of insufficient sample size and the limited time the initiative has been underway.

Participation in BPCI has continued to grow, with more providers entering Phase 2 in April and July 2015, and more EIs transitioning episodes to Phase 2. The increased sample sizes, as well as extended times under the initiative, will allow us to expand our understanding of the impact of BPCI and strengthen our conclusions about participation under Models 2 and 3. We will use this additional sample to make more distinctions across clinical episodes and understand the experience of subpopulations that may be more vulnerable to declines in quality. Participation in Model 4, however, has dropped, so there may be less that we can say about this model in next year's report.

I. Introduction

The Bundled Payments for Care Improvement (BPCI) initiative is designed to test whether linking the payments for all providers involved in delivering an episode of care can reduce Medicare costs while maintaining or improving quality of care. The Centers for Medicare & Medicaid Services (CMS) launched the BPCI initiative in 2013 under the authority of the Center for Medicare and Medicaid Innovation. BPCI Awardees, which can include hospitals, physician groups, post-acute care (PAC) providers and other entities, entered into agreements with CMS to be held accountable for total Medicare episode payments. Those agreements also specify Awardees' choices among four payment models, 48 clinical episodes, three episode lengths and waiver options.[7]

The BPCI initiative is designed to reward Awardees for adopting practices that reduce Medicare payments for the bundle of services in the episode relative to a target price, which CMS determines, based on the provider's historical payments for the same type of episode. For two of the four models, when Awardees' episode payments are below the target price, they may receive net payment reconciliation amounts (NPRA), which they can keep or share with their partnering providers. When Awardees' episode payments are above the target price, they may have to return amounts to CMS. Thus, in response to the BPCI incentive to attain positive NPRA, Awardees are expected to engage in behaviors to lower the cost of care, reduce the incidence of costly adverse events, care for lower cost patients, or some combination of these behaviors. These behaviors, in turn, can affect Medicare and beneficiary costs or quality of care in ways that may be desirable or undesirable.

The Lewin Group, with our partners, Abt Associates, Inc., GDIT, and Telligen, is under contract to CMS to evaluate and monitor the impact of BPCI Models 2, 3, and 4. This is the second of five Annual Reports that synthesizes the findings from various evaluation and monitoring activities under this contract.

A. BPCI Initiative

The BPCI initiative incorporates multiple approaches to aligning incentives for providers involved in an episode of care. Under each BPCI Model, an episode of care is triggered by a hospitalization for an MS-DRG contained in one of 48 clinical episodes (see **Appendix A** for a list of the 48 clinical episodes and MS-DRGs).[8]

The bundle is defined as the services provided during the episode that are linked for payment purposes. Certain services unrelated to the triggering hospitalization are excluded from the bundle, including readmissions for certain MS-DRGs and some Part B services. The bundle varies by Model as follows:

- **Model 2** has the most comprehensive bundle, which includes the triggering hospital stay (i.e., the anchor hospitalization), all concurrent professional services and post-discharge services, including hospital readmissions, delivered within the chosen episode length of 30, 60, or 90 days after discharge (with certain exclusions). Individual providers are paid

---

[7] This report describes the evaluation of Models 2-4; Model 1 is evaluated separately.
[8] **Appendix B** includes an acronym list and glossary for common terms used through this report.

on a fee-for-service basis and total episode payments are reconciled retrospectively against the established target price.

- The **Model 3** bundle includes services after the anchor hospital discharge, including professional services and readmissions within the chosen episode length of 30, 60, or 90 days (with certain exclusions). The episode starts when a beneficiary is admitted to a participating skilled nursing facility (SNF), home health agency (HHA), inpatient rehabilitation facility (IRF), or long-term care hospital (LTCH) following a hospitalization for a chosen clinical episode or when a beneficiary is admitted to a post-acute care (PAC) setting by a physician who is in a participating physician group practice (PGP). Individual providers are paid on a fee-for-service basis and total episode payments are reconciled retrospectively against the established target price.

- The **Model 4** bundle includes the anchor hospitalization, all concurrent professional services, and any readmissions and associated professional services that occur within 30 days of discharge that are not explicitly excluded from the bundle. Awardees are paid a prospectively determined amount and they, in turn, pay the providers involved in the episode. There is no NPRA for Model 4 because participants keep any difference between the prospectively determined amount and their episode costs.

There are almost 400 unique combinations of Model, clinical episodes, and participant type in BPCI. During the first year of the initiative, patient episodes were initiated in 103 of the possible combinations (See **Appendix C** for count of participants and patient episodes by Model and clinical episode during the first year of the initiative). In addition, for Models 2 and 3, Awardees may select one of three options for bundle length and risk track. Risk track refers to the winsorization, that is, the outliers that are excluded from the reconciliation payment calculation (1st to 99th; 5th to 95th; or 5th to 75th percentile).

CMS announced on August 23, 2011 that providers and other organizations could apply to participate for Phase 1, during which CMS and BPCI participants could prepare for the initiative. Phase 1 participants that were approved by CMS could then enter into an agreement with CMS and begin Phase 2, which involves the assumption of financial risk by Awardees. Awardees could enter into Phase 2 as early as October 1, 2013 for any of their chosen clinical episodes. By July 2015, all participants had to transition at least one clinical episode to Phase 2 to remain in the initiative. All participants must have transitioned all of their chosen clinical episodes to Phase 2 by October 2015, when Phase 1 ended.

1. Participant roles

There are several ways that organizations may participate in BPCI, distinguished by whether the participant is risk bearing, can initiate episodes under BPCI, or serves as an administrator or convener. An Awardee is a provider or other type of entity that has entered into the BPCI agreement with CMS and accepts risk. An episode-initiating (EI) provider may or may not be an Awardee, but it is the provider associated with the start of the episode. Under Model 2, an EI is a hospital or a PGP; under Model 3, it is a SNF, HHA, IRF, LTCH, or PGP; under Model 4 it is a

hospital. In this report, EIs and Awardees may also be referred to as participants. Additional terms and roles are described below.

- **Single Awardee (SA)** — Under Models 2, 3 and 4, SAs are individual Medicare providers that assume financial risk under the Model for episodes initiated at their institution.[9] These SAs are also Episode Initiators.

- **Awardee Convener (AC)** — Parent companies, health systems, or other organizations that assume financial risk under the Model for Medicare beneficiaries that initiate episodes at their respective Episode Initiating Bundled Payments Provider Organization (EI-BPPO) are ACs. An AC may or may not be a Medicare provider or initiate episodes.

- **Facilitator Convener (FC)** — An entity that submits a BPCI application and serves an administrative and technical assistance function on behalf of one or more Designated Awardees (DA) or Designated Awardee Conveners (DAC) is a Facilitator Convener (FC). Designated Awardees and DACs function as SAs and ACs, respectively, but join the initiative under a FC. Facilitator Conveners do not have an agreement with CMS, nor do they bear financial risk under the Model, or receive payment from CMS. The DA or DAC would have an agreement with CMS and assume financial risk under the Model for episodes initiated at their institution.

- **Episode Initiating Bundled Payments Provider Organization (EI-BPPO)** — Under Models 2, 3 and 4, EI-BPPOs are Medicare providers that deliver care to beneficiaries. EI-BPPOs are episode initiators associated with an AC or DAC. EI-BPPOs are where episodes begin under an AC or DAC. EI-BPPOs do not assume financial risk with CMS. They are associated with an AC or a DAC that assumes the financial risk.

- **Episode Initiators (EI)** — Under Model 2, an EI is the participating hospital where the BPCI episode begins or a participating physician group practice (PGP) if one of its physicians is the patient's admitting physician or surgeon for the anchor hospitalization. Under Model 3 an EI may be a participating PGP or a participating SNF, HHA, IRF, or LTCH that admits the patient within 30 days following a hospital discharge for a MS-DRG for the relevant clinical episodes (anchor hospitalization). Under Model 4 an EI is the participating hospital where the BPCI episode begins. Single Awardees and DAs are EIs. Awardee Conveners and DACs may or may not be EIs themselves, and also have one or more EIs under their Awardee structure.

2. *BPCI waiver options*

The design of the BPCI initiative allows Awardees to choose among several waivers of Medicare requirements to facilitate the implementation of care redesign interventions. An EI may or may not elect to use a waiver chosen by its Awardee.

---

[9] Under BPCI, assuming financial risk means that the entity would be obligated to repay the Medicare Trust Fund any Model 2 or 3 Net Payment Reconciliation Amounts (NPRA) or any Model 4 Reconciliation of Readmissions Amounts, and Excess Spending Amounts resulting from the Post Episode Spending Calculation.

Fraud and Abuse Waivers are owned by the HHS Office of Inspector General (OIG) and permit BPCI Awardees to engage in certain types of financial relationships that normally are not allowed.

- **Gainsharing** – A gainsharing waiver under Models 2, 3, or 4 allows BPCI participants to share incentive payments with gainsharing partners. Awardees must describe in their IP the specific methods for calculating and distributing these payments. The gainsharing partners may include an Awardee's EIs and other providers with a gainsharing agreement with the Awardee or the EI. Gainsharing is used to offer incentives to providers to support Awardees' care redesign initiatives. Awardees have many options for customizing their gainsharing methodology. Awardees can share savings generated internally, Internal Cost Savings (ICS), or NPRA, or both. Awardees may choose to share savings with individual physicians or other providers, determine when and how savings are calculated and distributed, and the manner in which the savings are contributed to various savings pools. The gainsharing calculation, which determines who receives incentive payments and how much they receive, may also differ across Awardees. Awardees can establish a fixed distribution schedule, or require gainsharers to meet specific efficiency, patient satisfaction, or cost savings metrics to qualify for distributions. Gainsharers must meet the quality metrics specified by the Awardee in its IP.

- **Beneficiary incentives** – With the beneficiary incentive waiver, an EI under any of the three Models may provide a service or product to a beneficiary that is related to the episode but not typically covered by Medicare. There must be a reasonable connection between the service or product and the beneficiary's medical care and the incentive must advance the beneficiary's clinical goal. Awardees must describe in their IP the criteria for beneficiary eligibility to receive the incentive as well as the clinical goal of the incentive.

Payment Policy Waivers are owned by CMS and permit BPCI Awardees to engage in certain behaviors that normally would not be covered (i.e. paid) by Medicare.

- **Three-day hospital stay waiver** – In general, Medicare beneficiaries are not eligible for Medicare-covered SNF care unless they have been a hospital inpatient for at least three consecutive days (not including the day of discharge) within 30 days of the SNF admission. Under this BPCI waiver, available only under Model 2, the SNF-qualifying hospital admission can be shorter than three days, as deemed appropriate by the treating clinicians. As a condition of this waiver, the majority of an Awardee's partner network must consist of SNFs rated three stars or better under the five-star quality rating system of Nursing Home Compare for at least 7 out of the 12 months immediately preceding a month in which the start of the Performance Year begins. In the IP, Awardees must describe how it plans to use the waiver, criteria for targeting beneficiaries for changes in care, the guidelines that will apply to discharging beneficiaries to SNFs prior to completing the three-day inpatient hospitalization, and how patient safety will be assessed while using this waiver.

- **Telehealth waiver** – Geographic restrictions on coverage of telehealth services furnished to Medicare beneficiaries may be waived for BPCI beneficiaries as long as the service is furnished consistent with other coverage and payment criteria.

- **Post-discharge home visit waiver** – The direct supervision requirement for home visits can be waived so that BPCI beneficiaries may receive a limited number of home visits (1

in a 30-day episode, 2 in a 60-day episode, and 3 in a 90-day episode) in the beneficiary's home by licensed clinical staff paid under the physician fee schedule.

3.   Purpose of the Annual Report

This Annual Report provides a summative and formative evaluation of the BPCI initiative based on the evaluation and monitoring activities that the Lewin team completed during the first two years of the contract. We summarize the participants' various approaches to implementing BPCI, including what influenced their selection of Model and clinical episodes, relationships with partners, use of waivers, and care redesign and cost saving strategies. We provide statistical analyses of changes in key outcomes from before BPCI to after implementation for BPCI episodes and comparable episodes from comparison providers. This report also provides data from the beneficiary survey on satisfaction, health care experience, and changes in functional status for various subgroups of beneficiaries in BPCI episodes relative to a comparison group. We also introduce additional analyses to address questions about market area effects of BPCI and factors that distinguish BPCI participants that achieved favorable payment and quality outcomes from those that did not. The quantitative analyses are based on the experience of Phase 2 participants during the first year of the initiative (episodes initiated between October 2013 and September 2014). The qualitative results reflect participants' experiences through June 2015. Please note that throughout this report, unless otherwise specified, the terms participant will refer to those in Phase 2.

B.   Research Questions

Three major evaluation and monitoring questions provide the framework for our analytic approach. Each question is addressed separately for the three BPCI Models under this evaluation. Under each major question are more detailed research questions that are addressed in this Annual Report.

   A.   *What are the characteristics of the BPCI initiative and participants at baseline and how have they changed during the course of the initiative?*

   B.   *What is the impact of the BPCI initiative on Medicare payments and the quality of care for Medicare beneficiaries?*

   C.   *What program, provider, beneficiary, and environmental factors contributed to the various results of the BPCI initiative?*

1.   What are the characteristics of the BPCI initiative and participants at baseline and how have they changed during the course of the initiative?

To understand initiative participants, their care redesign, Model incentive structures, and initiative adherence, we compiled data submitted by Awardees in their IPs and quarterly data submissions. We also used claims and patient assessment data to examine BPCI beneficiary characteristics. This information was supplemented by case studies, focus groups, and quarterly interviews with BPCI participants. We developed market profiles with data from the Provider of Service (POS), Area Health Resource Files (AHRF), and other secondary sources.   This information provided context and explanatory variables to understand the impact of BPCI and the factors that contributed to the results of BPCI. We also capture challenges implementing BPCI

shared during site visits, Awardee quarterly interviews, and interviews with episode initiators that had terminated their participation in BPCI.

- **Participants**—We describe BPCI participants to understand characteristics that may affect their ability or willingness to participate in this initiative, such as size, teaching status, profit status, market dominance, and proportion of patients that are Medicare beneficiaries. We also document characteristics that we hypothesize may affect their ability to redesign processes of care, negotiate with potential provider partners, control costs, or influence admitting or ordering physicians.

- **Market characteristics**—The structure of the health care market of the EIs may affect their ability to develop relationships with other providers or partnerships to deliver care across the entire bundle more efficiently. The market may also affect the care redesign opportunities for BPCI participants. We compare characteristics of markets where BPCI participants are located to characteristics of markets without a BPCI participant. Such characteristics include the overall competitiveness among providers, the availability of various types of providers, and Medicare managed care penetration.

- **Model incentive structure characteristics**—We document the choice of Model and episodes; the structure of the Awardee/EI (i.e. FC, AC presence); the level of risk; the role of the Awardee/AC/FC; and use of waivers for each Awardee as well as any changes in these during the course of the initiative. These BPCI design features may affect participants' ability to achieve program objectives.

- **Care redesign and cost saving strategy characteristics**—BPCI is designed to provide incentives to deliver care more efficiently while maintaining or improving quality. Awardees can achieve these objectives through care redesign or cost saving strategies. Awardees must document these strategies in their IPs and any changes have to be accepted by CMS. Awardees may submit changes to their IPs on a quarterly basis throughout the duration of their period of performance. We supplement the information from Awardee IPs with data from the Awardee interviews and site visits.

2. *What is the impact of the BPCI initiative on Medicare payments and the quality of care for Medicare beneficiaries?*

The Annual Report provides insights into the impact of BPCI on the costs of episodes, utilization of services, quality of care for Medicare beneficiaries, and provider referrals and market share. Our evaluation is designed to measure providers' responses to the incentives provided through participation in BPCI and how those behaviors affect the Medicare and beneficiary costs and quality of care.

- **Impact on payment and utilization**—Under BPCI, providers are expected to adopt care redesign and cost saving strategies that will change the use of health care services to reduce the cost of care in the bundle. We examined changes in the costs of care using

Medicare standardized allowed payments.[10] We accounted for differences in patient need for services through risk adjustment.

- **Impact on beneficiary quality of care—**We used several risk-adjusted measures to examine the impact of BPCI on the quality of beneficiary care, including mortality, readmissions, functional status, and clinical indicators. We also assessed the impact of BPCI on beneficiaries' experiences with care and improvements in functional status through the beneficiary survey.

- **Provider referral and market share -** We conducted an exploratory analysis to determine whether the BPCI initiative affects non-BPCI hospitals or PAC providers located in BPCI EIs' markets. We hypothesized that Model 2 and 3 EIs would leverage existing relationships with other types of providers to better manage their patients' care across multiple settings, and concentrate their patients' use of these preferred providers to the exclusion of others. We further hypothesized that Model 2, 3, and 4 BPCI EIs would strive to increase their market share of BPCI-eligible episodes, attracting patients that would otherwise go to competitors not participating in BPCI, to increase efficiency, savings, and profits. We used a PAC referral concentration index to examine whether the BPCI program induced changes in providers' referral patterns for PAC and also the market share of EIs to determine whether EIs have captured a greater market share of BPCI-eligible episodes over time.

3. *What program, provider, beneficiary, and environmental factors contributed to the various results of the BPCI initiative?*

For a selected set of clinical episodes with a significant number of EIs, we synthesized qualitative and quantitative findings to identify Awardee characteristics, care redesign approaches and related market, provider, and beneficiary characteristics that contribute to BPCI program success or failure. We linked outcomes to characteristics of the Awardees and the specific episodes they target to understand how these features of the initiative contribute to its stated goals of reducing costs while maintaining or improving the quality of care. We recognized that the interplay among multiple factors contributes to whether an Awardee reduces Medicare payments while maintaining or improving quality under BPCI. The analysis was conducted among the BPCI participants and their patients to determine what characteristics were associated with specific outcomes of interest evaluated under question B.

---

[10] These amounts combine the Medicare payments with the patient coinsurance and copayment amounts and then adjust for Medicare payment policies to ensure that any differences across time and providers reflect real differences in resource use rather than Medicare payment policies (e.g. teaching payments or differential payment updates). Our analyses consider payments that occur up to 60 days beyond the end of the episode.

II. Methods

The BPCI Models 2-4 evaluation and monitoring activities used a combination of quantitative and qualitative data and methods. The quantitative analysis used claims, assessment, and survey data to investigate the impact of BPCI on care experiences, clinical and functional outcomes, utilization, and costs. To this end, we rely on a non-experimental design, which used providers and their episodes in a comparison group to infer the characteristics of BPCI providers and episodes if there had been no BPCI initiative. We used qualitative data collected through case study site visits, focus groups, exit interviews, Awardee interviews, technical expert panels (TEP), and awardee-submitted data to examine how Awardees implemented their models, their success, and the factors that influenced the process and the resources that were required.

The discussion below outlines the main elements of our mixed-method approach including the data sources, identification of the study and comparison populations, outcome definitions, and the quantitative and qualitative analytical approaches.

A. Data Sources

The following section describes each of the quantitative and qualitative data sources used for the evaluation. Each data source addressed different evaluation needs including describing participants and their markets, building the comparison group, examining outcomes, assessing implementation, and testing factors associated with outcomes (See Exhibit 1). The mixed-method design required collecting primary data through a beneficiary survey as well as qualitative case study site visits, focus groups, and exit interviews; quarterly interviews with BPCI Awardees; a TEP; and Awardee-submitted data. Secondary sources included Medicare claims and enrollment, provider, and patient assessment data; Implementation Protocols; and publicly available data such as the AHRF.

Exhibit 1: Primary and Secondary Data Sources used in BPCI Evaluation and Monitoring Activities

Use in Evaluation	Primary Data	Secondary Data
Provide context to the BPCI initiative: BPCI participants, BPCI markets, care redesign approaches, and partnerships	▪ Case study site visits ▪ Focus groups ▪ Expert interviews ▪ Awardee interviews ▪ Awardee-submitted data	▪ CMS's BPCI database of BPCI Participant and Episode Reports, available via Salesforce ('CMS's BPCI database') ▪ Medicare Provider Enrollment, Chain, and Ownership System (PECOS) files ▪ Provider of Services (POS) files ▪ Area Health Resource Files (AHRF) ▪ Implementation Protocols
Assess implementation strategies and challenges	▪ Case study site visits ▪ Focus groups ▪ Expert interviews ▪ Awardee interviews	

Use in Evaluation	Primary Data	Secondary Data
Construct BPCI population and comparison groups	■ Case study site visits ■ Expert interviews ■ Awardee interviews	■ CMS's BPCI database , BPCI Participant and Episode Reports ■ Medicare Provider Enrollment, Chain, and Ownership System (PECOS) files ■ Provider of Services (POS) files ■ Area Health Resource Files (AHRF) ■ Master Data Management (MDM) ■ Medicare Part A and B claims ■ The Master Beneficiary Summary File ■ Episode SAS files from CMS' BPCI Reconciliation contractor
Evaluate BPCI impact on health and functional outcomes, utilization, and payments	■ Beneficiary survey	■ Medicare Part A and B claims ■ The Master Beneficiary Summary File ■ Minimum Data Set (MDS), Outcome and Assessment Information Set (OASIS), and Inpatient Rehabilitation Facility Patient Assessment Instrument (IRF-PAI) patient assessments
Evaluate patient-centered outcomes	■ Beneficiary survey ■ Technical expert panel	
Evaluate factors explaining differential outcomes	■ Awardee-submitted data	■ Medicare Provider Enrollment, Chain, and Ownership System (PECOS) files ■ Provider of Services (POS) files ■ Area Health Resource Files (AHRF) ■ Medicare Part A and B claims ■ The Master Beneficiary Summary File ■ Implementation Protocols

1. Secondary data

Exhibit 2 lists the secondary data sources and their uses for this study. We used provider-level data sources to identify BPCI participant providers, select valid comparison providers, as well as to describe both the study populations and the total provider population. Beneficiary-level Medicare claims and enrollment data were used to identify and construct episodes of care for patients at BPCI-participating sites (BPCI population) and at matched comparison providers. We also used beneficiary-level claims and patient assessment data to create outcome measures and beneficiary risk factors associated with the outcomes. Medicare claims data include claims for services incurred October 1, 2010 through December 31, 2014 and processed as of March 1, 2015. Patient assessment data include assessments completed January 1, 2011 through September 30, 2014 and processed as of March 1, 2015.

Exhibit 2: Secondary Data Sources used in BPCI Evaluation and Monitoring Activities

Dataset Name	Date Range	Dataset Contents	Use
Provider-level data sources			
CMS's BPCI database - BPCI Participant and Episode Reports	2013-2015	Information compiled by CMS on BPCI participants and potential future participants and their clinical episodes, including participant name, CMS Certification Number, location, type (ACH, SNF, etc.), BPCI "role", clinical episode type(s) and length(s), BPCI participation start and end dates, and contact information.	Used to identify Quarter 4 (Q4) 2013 through Quarter 3 (Q3) 2014 BPCI participating providers and clinical episodes. Identified potential future participants to exclude from comparison group.
Medicare Provider Enrollment, Chain, and Ownership System (PECOS)	2011-2014	Information on Medicare providers, including ownership and chain relationships among providers.	Used to identify ownership of BPCI providers and potential comparison providers and to create an indicator of whether the provider was part of a chain. Both of these characteristics were used in the creation of the comparison groups.
Provider of Services (POS) file	2011-2014	Information on Medicare-approved institutional providers, including provider number, size, and staffing.	Used within descriptive analysis of BPCI and non-BPCI participants as well as used to create predictors in provider propensity model on participation in BPCI or characteristics for Mahalonobis matching.
Area Health Resource File (AHRF)	2011	County-level data on population, environment, geography, health care facilities, and health care professionals.	Descriptive analysis of BPCI and non-BPCI market characteristics. Predictors in provider propensity model on participation in BPCI or characteristics for Mahalonobis matching.
Implementation Protocols	2013-2015	Information provided by an Awardee to CMS when joining BPCI (updated quarterly with any changes). The Awardee describes their care redesign activities, whether or not they will be participating in the OIG or CMS waivers, SNF partner lists, and the details of the beneficiary incentives and gainsharing methodology, if relevant.	Used to identify the count and percentage of Awardees and EIs participating in various care redesign and waivers. Used as potential characteristics of interest when evaluating what BPCI characteristics are associated with success/failure in the program.
Master Data Management (MDM)	2013-2015	Provider- and beneficiary- level information on participation in CMMI payment demonstration programs.	Used to identify providers who are involved in other ACO and Medicare Share Savings programs.
Episode files from Reconciliation contractor	2013-2014	Final episode SAS research dataset samples shared, when necessary, with the Reconciliation contractor.	Used to validate our implementation of the BPCI episode construction methodology.

Dataset Name	Date Range	Dataset Contents	Use
Beneficiary-level data sources			
Medicare Claims	Jan 2010-Dec 2014	Medicare Part A and B claims.	Used to create episodes of care and outcome measures such as readmissions, emergency department (ED) visits, number of days in each setting (e.g., acute care hospital, home health agency (HHA), skilled nursing facility (SNF). Also used to create risk factors including Hierarchical Condition Categories (HCCs) and health care use prior to anchor hospitalization.
Medicare standardized payments	Jan 2011-Dec 2014	Medicare standardized payments for 100% of Part A and B claims received via the IDR from another CMS contractor.	Used to create Medicare allowed standardized payment outcomes.
The Master Beneficiary Summary File (MBSF)	Jan 2010-Dec 2014	Beneficiary and enrollment information, including beneficiary unique identifier, address, date of birth/death, sex, race, age, and Medicare enrollment status.	Used to identify eligibility for episodes of care, beneficiary demographic characteristics, and beneficiary eligibility for inclusion in the denominator for each of the outcome measures.
Minimum Data Set (MDS) patient assessments	2011-2014	Comprehensive post-acute patient assessments completed by clinicians. Required for residents of Medicare-certified SNF facilities. Administered at entry to the facility, at discharge, and days 14, 30, 60, 90, and quarterly thereafter.	Provided functional status outcomes (early-, mid-, and long-form activities of daily living) for BPCI and comparison groups.
Outcome and Assessment Information Set (OASIS) patient assessments	2011-2014	Comprehensive post-acute patient assessments completed by clinicians. Required for Medicare-paid home health patients. Completed at the start of care and at discharge, and when care resumes following a hospitalization. Modified assessments are completed at recertification (60 days), if the patient's condition changes significantly, at transfer to an inpatient facility, and at death.	Provided functional status outcomes (bathing, upper- and lower-body dressing, ambulation/locomotion, and bed transferring) for BPCI and comparison groups.
Inpatient Rehabilitation Facility Patient Assessment Instrument (IRF-PAI)	2011-2014	Comprehensive post-acute patient assessments completed by clinicians. Required for all Medicare Part A fee-for-service patients who receive care from an IRF at admission and upon discharge. (For patients with a stay of less than 3 days, the discharge assessment was not required).	Provided functional status outcomes (self-care and mobility) for BPCI and comparison groups.

2. Beneficiary survey

The objective of the BPCI beneficiary survey was to explore any differences in patient care experiences and functional outcomes between Medicare beneficiaries cared for by BPCI providers and beneficiaries much like them whose providers are not participating in BPCI. The beneficiary survey collected information on a set of patient outcomes that were not available from other data

sources (e.g., provider communication) or that were not available for patients in all care settings (e.g., functional status measures.) In this report, we present results from the first two waves of BPCI beneficiary surveys, which were conducted in summer 2014 (Wave 1) and fall 2014 (Wave 2). This section describes the instrument, sampling, and administration of the beneficiary survey from which the results were based.

a. Survey instrument

The survey instrument was based on validated survey instruments, such as the CARE Tool, National Health Interview Survey, and Short Form 36 Health Survey (Gage et al., 2012; CDC, 2012; Brazier et al., 1992). New questions underwent cognitive testing with a small convenience sample of Medicare beneficiaries with recent hospital and PAC experience. The survey instrument used in Wave 1 was revised for Wave 2. Several Wave 1 survey items were modified in order to better assess patient independence in functioning[11] and three questions were added for Wave 2. Surveys from both waves are included in **Appendix D**.

The Wave 2 survey contained 36 multiple-choice, closed-ended questions and was designed to take an average of 25 minutes to complete. Survey questions covered a range of domains including functional status, overall mental and physical health, health care experience, and personal characteristics. (See Exhibit 3). For each of the seven functional areas, respondents were asked to recall their functional status before the anchor hospitalization, and also to report their current functional status at the time they were completing the survey, which was more than three months after the anchor hospitalization that starts a Model 2 episode or PAC initiation that starts a Model 3 episode.

Exhibit 3: Domain and Survey Items for Beneficiary Survey, Wave 2

Domain	Description
Functional Status	1) bathing/dressing/toileting/eating 2) planning regular tasks 3) use of a mobility device 4) walking by self without resting 5) walking up or down 12 stairs 6) physical or emotional problems that interfere with social activities 7) pain that interferes with normal activities
Overall mental and physical health status questions	1) how often respondent bothered by little interest in doing things 2) how often the respondent was bothered by feeling down, depressed or hopeless 3) overall physical health 4) overall mental health

---

[11] For Wave 2, questions in the Wave 1 survey that asked about the respondent's ability to walk across a room in a set period of time were revised to ask how far the respondent can walk without resting. Questions about walking up stairs were revised to specify a number of stairs, both up and down. One question about physical and emotional problems and one question about pain were revised as two questions referring to the period before the anchor hospitalization, and the time of survey completion.

Domain	Description
Health care experience	1) frequency of conflicting medical advice from medical staff 2) appropriate level of services received 3) frequency with which medical staff addressed the respondent in his/her preferred language 4) respondent feels that she/he was discharged at the right time 5) medical staff took patient preferences into account when arranging for health care services after discharge 6) respondent had a good understanding of how to take care of herself or himself prior to discharge 7) medical staff clearly explained how to take medications; 8) medical staff clearly explained needed follow-up appointments 9) respondent and caregivers ability to manage their health care needs 10) overall satisfaction with recovery since discharge
Personal characteristics	1) lives alone, with others, or with paid helper 2) gender 3) education level 4) ethnicity 5) race

b. Survey strata

The beneficiary survey uses a stratified sampling method with matched BPCI and comparison group beneficiaries within each cell defined under each stratum. The number of strata varied by wave based primarily on the number of strata for which sufficient BPCI episodes were available to ensure at least 310 BPCI responses, given expected response rates (see Power calculation and response rate assumptions below). We conducted strata selection within Model 2 and 3 separately.[12] Each stratum selected has two arms: the BPCI group and a matched comparison group.

The Wave 1 survey had four sampling strata:

- Model 2 episodes with a major complicating condition (Model 2, MCC)

- Model 2 episodes without MCC

- Model 3 episodes with MCC

- Model 3 episodes without MCC

The Wave 2 survey had four sampling strata: [13]

- Model 2 "major joint replacement of the lower extremity" (MJRLE) episodes

- Model 2 "non-surgical: cardiovascular" episodes [14]

---

[12] We have not surveyed any patients from Model 4 since the number of episodes within Model 4 was not adequate to meet the minimum sample size of even a single stratum based on power analysis.

[13] We changed the sampling strata from Wave 1 to Wave 2 to better understand the impact of the BPCI program on specific clinical episode groups.

[14] This clinical episode group includes the following episode groups: acute myocardial infarction, cardiac arrhythmia, atherosclerosis, chest pain, medical peripheral vascular disorders, syncope & collapse, and congestive heart failure.

- Model 2 "non-surgical: respiratory" episodes [15]
- A single stratum including all Model 3 episodes

Within each stratum, BPCI and comparison beneficiaries were matched within cells, which were defined based on beneficiary age groups, BPCI clinical episode groups and provider type (e.g. provider size and provider academic status). There were 148 final sampling strata in Wave 1 and 174 in Wave 2, which is a combination of the four main sampling strata, two groups (BPCI and comparison), and the sampling cells within each stratum.

c. Survey administration

We mailed each sampled beneficiary a paper survey, and several reminders and re-mailings, and then followed-up by telephone with those for whom a phone number was available. The first survey was mailed to beneficiaries within about 90 days after their hospital discharge for Model 2 and within about 120 after episode initiation for Model 3.

For Wave 1, we mailed surveys to a total of 5,947 beneficiaries (BPCI and comparison groups combined) and 3,030 surveys were returned with at least one question answered. The overall response rate in Wave 1 was 50% for the BPCI group and 52% for the comparison group. In Wave 2, we mailed surveys to a total of 6,162 beneficiaries (BPCI and comparison groups combined) and 3,008 surveys were returned with at least one question answered. The overall response rate in Wave 2 was approximately 49% for both the BPCI and comparison groups.

In both waves, response rates varied considerably across sample strata. Generally, response rates were higher for patients in Model 2 than for those in Model 3. In Wave 1, within each model, response rates for patients with MCC were lower than for patients without MCC. In Wave 2, for Model 2, the response rates were higher among patients with major joint replacement of lower extremity episodes than for those with non-surgical cardiovascular and respiratory episodes. Within Model 3, response rates were higher in episodes initiated by a HHA than those initiated by a SNF. These results suggest that beneficiaries who were more seriously ill, or required institutional care after their hospitalization, were less likely to respond to the survey – possibly because they had not yet returned home and did not receive the survey. In both waves, there were no statistically significant differences in response rate between BPCI and comparison groups in the various strata.

d. Power calculation and response rate assumptions

We tested the null hypothesis that the population percentage of a binary response in the BPCI sample is equal to the percentage in the comparison sample. That is, if the BPCI sample has a functional improvement rate of X%, and the comparison sample has a functional improvement rate of Y%, we can be confident that X and Y are different due to the effect of BPCI rather than random chance. We determined that a combined target sample size of 620 completed surveys (310 each for the BPCI and comparison groups, per stratum per wave) would enable us to reject the hypothesis of no difference in population percentages of our outcomes of interest with 80%

---

[15] This clinical episode group includes the following episode groups: chronic obstructive pulmonary disease, bronchitis, asthma; other respiratory; and simple pneumonia and respiratory infections.

power when there is a true underlying difference of 10 percentage points in a binary variable with a baseline value of 50%.

We used estimated response rates to determine the size of the initial sample required to yield 310 completed surveys in each group. For Wave 1, the first BPCI beneficiary survey, we estimated a response rate based on recent surveys our team had conducted with Medicare beneficiaries with varying degrees of functional impairment and disability. We assumed a survey response rate of 49% for Model 2 patients and a lower survey response rate of 36% for Model 3 patients. For Wave 2, we started with the actual response rates from Wave 1 and subtracted the margin of error from these response rates to reach a conservative estimate. For example, the actual response rate of respondents with "major joint replacement of the lower extremity" in Wave 1 was 74% with a margin of error of 5%. So we used a response rate of 69% (74%-5%) for respondents with "major joint replacement of the lower extremity" in Wave 2.

3. Case study site visits

Case studies were based on two-day, in-person site visits that involved interviews with key individuals responsible for different aspects of BPCI implementation and management, including clinical and administrative leaders and operational staff, at episode-initiating sites (both Awardees and EIs under Awardees). The information collected during each site visit complements data submitted by Awardees through their Implementation Protocol and quarterly data submissions. Most case studies incorporate interviews with leaders and operational staff from the episode-initiating sites. Some site visits have included interviews with staff from the Awardees or Conveners if applicable.

During our site visits conducted during Q4 2013 through Quarter 2 (Q2) 2015 we focused on why organizations chose to participate, how they selected their clinical episodes and their partners, their initial infrastructure investments to participate in BPCI, and their goals for the BPCI initiative. We also asked about the processes they adopted to meet the incentives offered through participation in BPCI, including those used in contracting, gainsharing, care redesign, quality and cost monitoring, and their implementation experiences, both the challenges and perceived successes. We continued to ask these questions during the second year of the initiative and added questions on new topics, including: reconciliation results (for sites that completed at least one reconciliation cycle), implementation of beneficiary incentives, and status of enacting the care redesign activities described in Implementation Protocols.

a. Case study sites

Case study sites were selected based on descriptive characteristics that inform a wide range of BPCI approaches and perspectives. The study sites varied in several key aspects that could affect provider incentives and the impact of the intervention. These aspects include:

1. Model: Each Model incentivized a different group of providers. While the same types of providers may be involved in patient care, only the targeted ones in each Model may be affected by the incentives designed into the model (e.g., while hospitals are part of a patient's episode in Model 3, the bundle excludes those services from the cost and outcome estimates).

2. Awardee Size: The number of EIs under an Awardee could be related to their outcomes. A large number of EIs could mean more opportunities for the organization to learn,

thereby enhancing their ability to develop effective care redesign processes. Single Awardees may not have the same ability to benefit from shared learning.

3. Convener approach: We are interested in understanding the impact of a convener on success under BPCI and the various services and information exchanged between conveners, awardees, and EIs.

4. Tenure in the initiative: Earlier BPCI entrants were selected for a greater share of the site visits than later entrants in the second year of the contract because they have had more time in the initiative and therefore more experience and knowledge to share during the site visit.

Exhibit 4 displays the 23 selected case study sites during Q4 2013 through Q2 2015 on a national map with all BPCI participants. The selected sites included 13 hospitals, six SNFs, 3 PGPs, and 1 IRF. Fifteen sites had an Awardee Convener, five had a Facilitator Convener, and three were Single Awardees.

Exhibit 4: Geographic Location of Case Study Participants among all Q4 2013 - Q3 2014 BPCI Episode Initiators
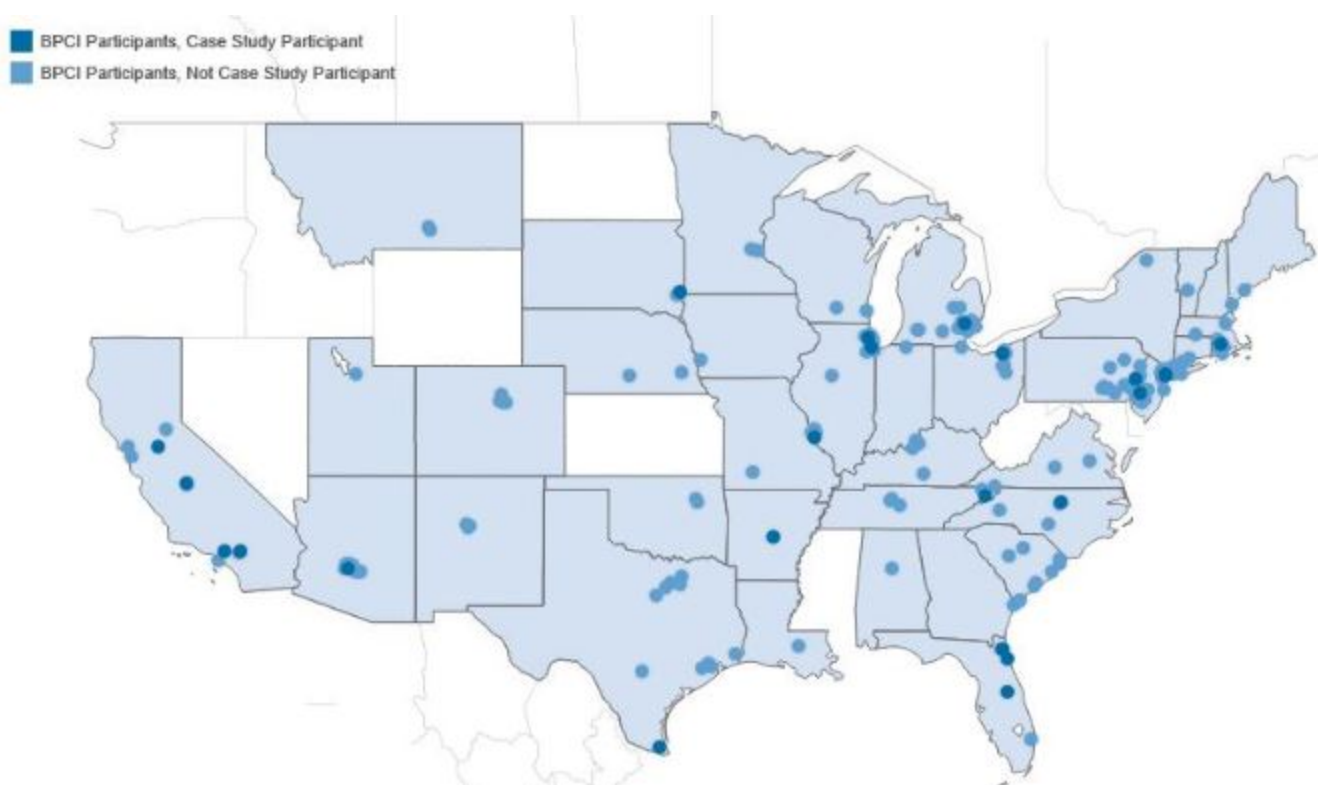
Source: Lewin analysis of CMS's BPCI database, as of May 2015, on BPCI participants from Q4 2013 through Q3 2014 and 2013 POS data.

NOTE: Data points may overlap in locations with multiple participant EIs. Two additional case study participants starting in Year 2 (January 2015) have been added to this map for complete representation of case studies in this report.

Exhibit 5 compares the characteristics of the 23 case study sites to all Q4 2013 through Q2 2015 BPCI EIs. As the evaluation continues, future sites will be selected to maintain a case study sample that is representative of all BPCI EIs.

Exhibit 5: Characteristics of Case Study Participants Q4 2013 to Q2 2015 and all BPCI Episode Initiators

	Case study participants (N=23)		All BPCI episode initiators in Q4 2013 through Q2 2015 (N=678)	
Model				
2	11	48%	349	51%
3	9	39%	309	46%
4	3	13%	20	3%
Participant Role				
Designated Awardee	3	13%	98	14%
Single Awardee	3	13%	31	5%
Episode Initiators (that are not Awardees)	16	70%	544	80%
Awardee Convener	1	4%	5	1%
Type of Participant				
Skilled Nursing Facility	6	26%	248	37%
Home Health Agency	0	-	51	8%
Acute Care Hospital	12	52%	272	40%
Physician Group Practice	3	13%	100	15%
Inpatient Rehabilitation Facility; Long-term Care Hospital	2	9%	7	1%
Clinical Episodes				
1) Orthopedic surgery	21	91%	499	74%
2) Non-surgical: other medical	9	39%	285	42%
3) Non-surgical: neurovascular	5	22%	189	28%
4) Non-surgical: respiratory	10	43%	359	53%
5) Non-surgical: cardiovascular	11	48%	347	51%
6) Non-surgical and surgical: GI	6	26%	187	28%
7) Cardiovascular surgery	9	39%	194	29%
8) Non-surgical: Ortho	8	35%	199	29%
9) Spinal surgery	9	39%	146	22%
Geographic Region				
Northeast	5	22%	178	26%
South	7	30%	214	32%
West	5	22%	122	18%
Midwest	6	26%	164	24%

Source: Lewin analysis of CMS's BPCI database, as of July 2015, on BPCI participants from Q4 2013 through Q2 2015.

b. *Interview protocols*

The site visit protocols were designed to gather information about the design, implementation, and initial results of BPCI from EI clinical and administrative leadership and managers involved with the initiative. Questions pertained to BPCI entry decisions and structure, experience with

BPCI, market effects, successes and challenges, ability to replicate, quality management, care redesign, and care management.

Protocols for case studies are attached in **Appendix E**. Separate interview protocols were tailored to and used for each type of respondent, ensuring consistency in question presentation and that the questions were appropriate for the interviewee.

During case study site visits BPCI leadership was asked about decisions that led to joining the initiative and why they chose to participate. They were asked about their partners (e.g., post-acute care providers, physician groups, etc.), care redesign approaches, gainsharing, and why they chose their options for each of these topics. They were also asked how they will determine whether their approaches are successful and what they expect to gain.

Operational managers were also interviewed, including financial managers, clinical managers, quality and outcomes directors, and data and IT managers involved in the BPCI initiative in each site. To capture the experience of staff from all levels of the organization, interviews were also conducted with clinical staff (e.g., case managers, nurses, and therapists) who provide care directly to BPCI beneficiaries. Interviewees were identified in consultation with leadership at each site through several planning calls in advance of the site visit. Convener staff were included where appropriate, although all case study site visits focused on the experience of the episode initiator. Interviewees were asked about their expected goals for their tasks related to the initiative, how their efforts differ from prior practice in their organizations, how their jobs have changed, the types of materials or practice programs they put in place to effect changes, and why the approaches were chosen. They were also asked about their perceptions about actual implementation and whether they viewed the initiative as meeting its stated goals. Interviews typically lasted one hour with each respondent.

4. Focus groups

In addition to the case study site visits, qualitative data were also collected through two focus groups on the following topics: 1) care redesign across EIs under the same Awardee Convener and 2) relationships between EIs and PAC partners. These focus groups complemented the site visit data and expanded our understanding of the effect of BPCI on participants, their partners and their markets. Focus group participants were staff members with sufficient experience on these topics who were able to offer personal insights, experiences, and opinions to the interviewing team. Participants included care coordinators, case managers, and nurses with comparable levels of responsibility. We selected participants with the same level of seniority, to the extent possible, because we wanted them to feel comfortable in expressing their opinions.

Focus groups generate discussion that elicits insights from knowledgeable actors that would not surface during the regular interviews conducted during the case study site visits. Ideally, participants in the group should benefit from the discussion and insights offered by others. Both focus groups included individuals who would not regularly gather and discuss their shared experiences.

Focus groups were conducted for one Model 2 participant and one Model 3 participant. Information on the topic and number of participants is included in Exhibit 6 below. Focus group sites were selected based on sites' ability to identify a sufficient number of staff with the experience needed to

yield a successful focus group. For example, when identifying possible sites for the "care redesign under the same convener" focus group, we determined that a convener should have staff from a minimum of three EIs to participate. This approach ensured that we had enough individual participants with the knowledge and experience to sustain discussion on these topics.

Exhibit 6: Focus Groups conducted in 2014

Topic	Participants	Topics
Care Redesign Under the Same Convener	Three Episode Initiators in Local Region Under Single Facilitator Convener	▪ Relationships between EIs and ACs ▪ AC services provided to EIs ▪ How ACs adapt their approach to address needs of the distinct EIs
Relationship Between PAC providers and EIs	Six PAC providers (Three SNFs and Three HHAs) in Local Region Working with Episode Initiator	▪ How EIs and PAC providers work together change processes for 1) care delivery, 2) hospital discharge and 3) PAC admission ▪ PAC providers' views of BPCI

The Lewin team, in consultation with CMS, selected the overall topics and questions for these focus groups. The protocols were designed to elicit important information for the evaluation while allowing the flexibility to follow the issues identified by participants. Questions were tailored to address the goal of the focus group; therefore, a unique protocol was created for each focus group topic. The protocols for each focus group topic were split into two sections. The first section included the main body of questions for the focus group topic and the second section addressed lessons learned. Scripted introductions were prepared to explain the goals of the focus group to participants. Focus groups were 90 minutes in length and were conducted in-person. The protocols used for these focus groups are included in **Appendix F.**

5.   *Expert interviews*

Expert interviews provided another source of qualitative data. Similar to the focus groups, expert interviews complemented the data collected during case study site visits and expanded our understanding of the effect of BPCI on participants, their partners, and their markets. In consultation with CMS, we have focused all of our expert interviews to date on participants that have terminated their participation in BPCI.

Upon receiving notice of their termination from CMS, we reached out to BPCI participants that withdrew from the initiative. A letter to these sites requested a 60-minute call with their key staff involved in the BPCI initiative to discuss their reasons for ending their participation. Nine of 15 organizations that were contacted participated in our requests for expert interviews (the 15 organizations are listed in Exhibit 7). Interviews were typically conducted with two to three individuals.

Exhibit 7: Organizations that ended Participation in BPCI and were
Contacted for an Interview

Model	Organization BPID	Organization Name
2	2054-001	Bayonne Medical Center
2	2077-001	Touchstone Health
2	2078-001	Vanguard Health Chicago
2	2900-000	SSM Managed Care Organization
2	2308-000	Pocono Medical Center
2	2314-000	Harrisburg Hospital, Community General Osteopathic Hospital, and The West Shore Hospital
2	2802-000	Billings Clinic
2	6245-000	West Houston Medical Center
2	6401-001 (4060-001 previously)	Summa Health Care System
3	3057-000	Amedisys
3	3403-001	Premier Health & Rehabilitation
4	4022-000	St. Luke's Hospital
4	4058-001	Sisters of Charity of Leavenworth Health System
4	4064-001	Abrazo Region Services
4	4067-000	University of Colorado Hospital Authority

Designed with input from CMS, the protocols for the expert interviews elicited information on potential challenges that may affect the ability to scale this initiative to a broader group of providers. These terminated EI respondents described various topics, including: key factors in the decision to withdraw from BPCI, ways that the program could be improved for future participants, and what the site could have done differently to have been successful in BPCI. Interviews lasted between 30 and 60 minutes. The full protocols for the expert interviews are included in **Appendix G**.

6. *Awardee interviews*

The Awardee interviews were semi-structured interviews lasting up to one hour with the Awardee's choice of representatives. We conducted Awardee interviews on a quarterly basis with the goal of interviewing 20 to 25 Awardees each quarter. Interviews were conducted over a period of two to three weeks prior to the end of the calendar quarter. Upon completion of all interviews in a given quarter, we analyzed the results and shared a summary of key findings with CMS. These findings informed other qualitative and quantitative data analyses.

In selecting the Awardees for a given quarter, we aimed to ensure that each quarter's sample had a comparable mix of Awardees and appropriately inform the domains that were being explored in a given quarter. Although the target was to conduct 20 to 25 interviews each quarter, the sample sizes varied depending on the topic that the quarterly interviews addressed. For example, we conducted our Quarter 1 (Q1) 2015 quarterly interviews with all FCs (n=8 as of January 2015)

as the protocol was specific to issues FCs alone may encounter. In this report, we summarize findings from the Awardee interviews conducted during the following quarters: Q1 2014, Q2 2014, Q3 2014, Q4 2014, and Q1 2015.[16] From Q1 2014 to Q1 2015, we conducted 85 quarterly interviews with a total of 77 Awardees and eight FCs. We interviewed all of the Awardees that started Q4 2013 and a portion of the Awardees that started Q1 2014. Exhibit 8 displays the model, role, and number of episode initiators among the sample of Awardees with whom we held interviews.

Exhibit 8: Characteristics of Awardees and Facilitator Conveners Interviewed, Q1 2014 - Q1 2015

		Q1 2014-Q2 2014 (N=37)		Q3 2014-Q1 2015 (N=48)[17]	
		N	**%**	**N**	**%**
Model	2	24	65	33	69
	3	10	27	8	17
	4	3	8	7	15
Role	DA	16	43	15	31
	AC	10	27	10	21
	SA	9	24	14	29
	DAC	2	5	1	2
	FC	0	0	8	17
Number of EIs	0	0	0	8	17
	1	25	67	29	60
	>1	11	30	8	17
	>10	1	3	3	6%

Source: Lewin analysis of CMS's BPCI database, for Phase 2 Awardees as of April 2015.

The Lewin team, in consultation with CMS, selected questions and topics, as informed through the qualitative and quantitative analyses. For Q1 and Q2 2014, we used similar protocols directed at the interviewee's decision to participate in the initiative, rationale for decisions on model characteristics (i.e., model, episodes, and bundle length), waiver use, and formal and informal partnerships. The same Awardee interview protocol was used for Q3 and Q4 2014, with minor changes made to improve the clarity and flow of the interviews. These quarterly interviews covered the Awardee's care redesign implementation and cost-saving strategies across its EIs. In Q1 2015, we designed and used a separate interview protocol for FCs. These interviews included questions on the rationale to participate as and the role of the FC in BPCI. Protocols for quarterly interviews are attached in **Appendix H.**

---

[16] Our analysis of Q2 2015 Awardee interviews will be included in OY2 Annual Report.

[17] Count represents the total number of Phase 2 Awardees in Salesforce at the time we had created our quarterly Awardee interview sample (May 2014).

7. Technical expert panels

Five TEPs planned over the course of the evaluation will gather clinician insights into patterns of care and changes in care for BPCI beneficiaries that are identified through the case studies or quantitative data analysis. The TEPs focus on a single clinical episode or diagnosis; therefore, the TEP panelists represent the range of clinicians and specialists that care for that type of case. The TEP members are charged with identifying inappropriate patterns of care, new methods of treating patients, and factors we should track to be able to adequately report on quality outcomes.

The first TEP focused on beneficiaries with major joint replacement of the hip or knee under all three BPCI Models. The objectives were to:

- Identify potential outcomes associated with the observed care patterns.

- Identify care patterns that signal questionable care.

- Identify markers of appropriate care and of potential inadequate care.

- Identify beneficiary populations that may be particularly susceptible to poor quality care.

- Identify measures to include in quarterly rapid cycle evaluation reports and other quantitative analyses of utilization and outcomes.

a. TEP panelists and administration

Eight panelists were identified through professional contacts and CMS suggestions and vetted by CMS. The panelists included two board certified orthopedic surgeons, a certified rehabilitation registered nurse, a home health physical therapist, an occupational therapy expert, a retired home health director, an internationally recognized geriatrician, and a PhD physical therapist who specializes in outpatient orthopedics with geriatric populations. See **Appendix I** for panelist names and the report of findings. Dr. Christine LaRocca, a board certified geriatrician and medical director for Telligen, facilitated the TEP.

The two hour TEP occurred via webinar on May 6, 2015. We created a pre-work packet consisting of TEP relevant BPCI background information, panelist biographies, an agenda, presentation slides, and general expectations for the TEP. CMS approved the materials prior to distribution to the panelists. The presentation slides included additional BPCI information as well as data and questions for the panel.

Through webinar technology, all participants viewed the same slides and were encouraged to submit comments or questions orally, or through online chat features. Dr. LaRocca ensured each panelist was given an opportunity to speak to each question, which was discussed for approximately 20 minutes. The entire meeting was recorded to ensure accurate recall of the discussion.

b. Topics and discussion

Experts were asked to comment on specific BPCI findings from the first two quarters of BPCI (episodes initiated between October 2013 and March 2014) to support interpretation of the data and develop new directions for analysis. Given the timing of the TEP, this preliminary evidence reflected only early joiners. (The full four quarters presented in the rest of this annual report were

not available yet.) Exhibit 9 presents a summary of the key topics that were discussed and the specific questions asked of the TEP panelists.

Exhibit 9: Key Topics and Questions for Technical Expert Panel on Major Joint Replacement of Hip and Knee

Topic / BPCI Finding	Questions to Panelists
SNF and HHA use: Relative to a comparison group, BPCI patients: ▪ were discharged less often to a SNF (Model 2). ▪ had shorter lengths of stay in SNF (Models 2 and 3). ▪ had increased use of HHA services (Model 2).	▪ Which patient populations may be particularly susceptible to suboptimal outcomes with these care patterns? ▪ What unintended consequences should we be aware of, and how might we measure them? ▪ What should we look for with respect to functional outcomes given the different capabilities of these settings?
Physical Therapy: No physical therapy after hip replacement (Model 2)*	▪ For which patients might this always or never result in a good outcome? ▪ What unintended consequences should we be aware of, and how might we measure them? ▪ What less invasive hip procedures, such as anterior hip replacement, might we see in the claims data now and in the near future? ▪ When were they developed and how frequently are they performed?
Preoperative programs: ▪ Patients are encouraged to participate in a "Pre-hab" exercise program (Model 2). ▪ Patients are required to attend mandatory total joint replacement education classes (Model 2).	▪ Taken individually or in combination, are these exercises and educational programs important contributors to high quality outcomes? ▪ Is the requirement to participate a subtle form of cherry picking?
ED visits: Increased ED visits without hospitalization within 30 days of discharge for BPCI patients (Model 2).	▪ It is possible that some of these visits are planned? ▪ What do you think about planned ED use as part of care redesign? ▪ How do you interpret this finding of increased ED visits?
SNF v. IRF recommendation	▪ What are the factors you consider when you recommend discharge to an IRF?

*This question was further clarified during the TEP to, "No outpatient physical therapy (PT) after hospital discharge for total hip replacement (THR)." Typical practice, not limited to BPCI participants, was generally described as 1-3 PT sessions while in the hospital after elective primary THRs, with the vast majority of patients not receiving ongoing outpatient PT.

8. Awardee-submitted data

Beginning Q1 2015, Awardees submitted data to Lewin on a quarterly basis so that we may: 1) track waiver use and assess waiver adherence; 2) measure quality with data not available through secondary sources; 3) document participant characteristics; and 4) gather initiative-related information, such as progress implementing care redesign. [18,]

BPCI participants submitted data for the first time in Q1 2015 based on activities that occurred from Q4 2013 to Q4 2014. Analysis of the data collected during this initial submission period is

---

[18] The gainsharing section of the Awardee quarterly data collection has been delayed. The remainder of this section refers to the structure and content of the quarterly data collection process, which will exclude gainsharing.

included in this report, while analysis of data collected during subsequent quarters will be included in future reports. We collected data directly from the Awardees through our online reporting platform. Awardees that had active Phase 2 episodes at any point between Q4 2013 and Q4 2014 were required to submit data. Data submission occurred at the Awardee level and each Awardee was responsible for ensuring that data was submitted for any affiliated EIs. As illustrated in Exhibit 10, there were 62 Model 2 Awardees, 20 Model 3 Awardees and 13 Model 4 Awardees that were required to submit data during the first submission period. Of these, 85% of Model 2, 95% of Model 3, and 85% of Model 4 Awardees submitted data.

Exhibit 10: Number of Participants Required to Submit Data and the Percentage of Participants who Submitted Data during the First Data Submission Period by Model, Q4 2013 - Q4 2014

	Model 2		Model 3		Model 4	
	Required (N)	Submitted N (%)	Required (N)	Submitted N (%)	Required (N)	Submitted N (%)
Single Awardee	16	15 (94)	1	1 (100)	7	7 (100)
Awardee Convener	16	13 (81)	7	7 (100)	4	2 (50)
Designated Awardee	28	23 (82)	11	10 (91)	2	2 (100)
Designated Awardee Convener	2	2 (100)	1	1 (100)	-	-
Total number of Awardees	62	53 (85)	20	19 (95)	13	11 (85)

Note: Facilitator Conveners are not required to submit data and are therefore not included in the exhibit.
Source: Lewin analysis of Awardee-submitted data collected February 2015 for Model 2, 3 and 4 episode initiators participating in BPCI between Q4 2013 – Q4 2014.

Exhibit 11 includes all of the data elements that we collected from the Awardees.

Exhibit 11: Awardee-provided Data Elements

Domain	Data element	Definition
Participant baseline characteristics	Patient Mix	The proportion of admissions and patient days by payer
	Prior Care Redesign Experience	Prior experience with care redesign initiatives
	Payment Incentives Experience	Prior experience with payment incentives
	EHR Use	Use of an electronic health record/electronic medical record system
	Meaningful Use Functionalities	Meaningful use functionalities that are available and used
	Health information exchange capabilities	Availability and use of health information exchange capabilities
	Health information exchange method	Method used to exchange information with beneficiaries and providers
BPCI-related activities	Status of Care Redesign	Progress in implementing care redesign activities

Domain	Data element	Definition
Quality monitoring measures	Medication Reconciliation at Discharge	The percentage of BPCI patient discharges for members 65 years of age and older for whom medications were reconciled at discharge from the hospital
	Medication Reconciliation at Admission or within 24 hours of Admission	The percentage of BPCI patient discharges for members 65 years of age and older for whom medications were reconciled at admission or within 24 hours of admission to a PAC facility
	Patient death or serious injuries reportable to FDA	Patient death or serious injury associated with the use of contaminated drugs, devices, or biologics provided by the health care setting and patient death or serious injury associated with the use or function of a device in patient care, in which the device is used or functions other than as intended
Beneficiary incentive waiver adherence and use	Beneficiary Incentive List	List of beneficiaries receiving incentives, the person/entity, administering the item/service, date the item/ service was provided, the date the item/service was received
	Incentives and Total Number of Beneficiaries Eligible to Receive Incentive	Description of any incentives offered and the total number of BPCI beneficiaries eligible to receive each incentive, based on beneficiary identification criteria in Implementation Protocol

B. Study Populations

In this section we describe the BPCI population and the methodology for constructing the comparison group. We also specify the BPCI population and comparison group, if any, used in each analysis.

1. BPCI population

The BPCI population included all episodes initiated by EIs that had Phase 2 episodes between Q4 2013 and Q3 2014. If an EI terminated during this time period, we included the episodes that they initiated up until their withdrawal date. Exhibit 12 indicates the Model and EI type included in each analysis of the evaluation. For Models with very few participating EIs in certain provider categories, low sample size prohibited some analyses.

Exhibit 12: Model and EI[19] Types Included in each analysis included in this report

	EI Type	Number of BPCI EIs, Q4 2013 – Q3 2014	Characteristics of the Program	Impact of BPCI (DiD estimates)	Beneficiary survey	Provider referral and market share	Factors Contributing to Differences across BPCI Providers
Model 2	Hospital	110	X	X	X	X	X
	PGP	3	X				
Model 3	SNF	63	X	X	X	X	X
	HHA	28	X	X	X	X	X
	IRF	1	X				X
	LTCH	1	X				
	PGP	1	X				
Model 4	Hospital	20	X	X		X	X

2. Selection of providers in comparison group

The difference-in-differences (DiD) approach requires a comparison group of non-BPCI providers ("non-participants") that were similar to the BPCI providers with respect to market, available services, and case-mix. Because providers voluntarily enroll in BPCI, BPCI participants were likely to be different than non-participants, which could affect patient outcomes. BPCI participants may have less efficient care and larger room for improvement relative to non-participants. This self-selection could result in a biased estimate of the impact of BPCI on outcomes. BPCI participants may have improved outcomes over time even without participating in BPCI. Moreover, program evaluation literature indicates that treatment effect estimates based on standard regression models can be very sensitive to untestable model assumptions when the intervention and comparison group are dissimilar in one or more dimensions (Dehejia and Wabha (2002), Zhao (2004), Smith and Todd (2005)).

During the first two years of the evaluation, we constructed comparison groups for each Model and provider type and clinical episode group, with the exception of PGP EIs, from the universe of Medicare providers that had not signed up for BPCI in either Phase 2 (active period of performance) or in Phase 1 (active preparatory period). Please note that the comparison group for the Model 3 participants was not a close match on several variables, particularly baseline episode costs. Therefore, any Model 3 results should be viewed with caution. In addition, the data linking individual physicians to PGPs was not always accurate and may have associated physicians with a BPCI-participating PGP who were not actually in that PGP. Similarly, certain physicians who were members of a BPCI-participating PGP may not have been linked to that PGP. The frequency of these attribution errors is not known. Thus, results about PGP-initiated episodes should be viewed with caution.

The comparison population for the claims analysis was selected in four steps. First, to try to account for potential selection bias, we compare Phase 2 BPCI participants to non-participants

---

[19] Throughout the report, EIs refer to any entity, including Awardees, that initiate episodes.

with the goal of selecting non-participants that were similar to BPCI participants. Based on this comparison, we created a set of exclusion rules to restrict the group of providers included in the comparison group. Second, each BPCI EI was matched to 15 comparison providers based on market and provider characteristics and several baseline claims outcomes using propensity score matching or Mahalanobis matching. We were not able to match on all key outcome measures, notably baseline mortality rates, because of the multiple outcomes and relatively small sample sizes. Mortality was particularly problematic because small sample sizes can lead to extreme values because mortality is a relatively rare event. (The matching method was chosen based on number of BPCI EIs participating in a given model/EI/clinical episode group).[20] Third, episodes were constructed for beneficiaries treated by BPCI and comparison providers who met the inclusion criteria. Fourth, a sample was drawn from among beneficiaries treated by comparison providers identified in the previous step to match the distribution of MS-DRGs and dates of service in the BPCI population.

Step 1: Exclude ineligible non-participating providers

We began by defining the set of providers that are eligible for inclusion in the comparison group. We applied exclusions for each Model, EI type, and clinical episode group separately based on market and provider strata. First, we excluded any providers that were not in the same ownership (i.e., government, non-profit, for-profit) or location (i.e., rural/urban) that were not represented in the BPCI population. Additionally, non-participating hospitals were excluded if they were located in markets with high BPCI participation, not paid under Medicare's inpatient prospective payment system, owned by a BPCI-participating organization, or preparing to join BPCI.

Market exclusions

We defined the market for a given provider as its Core-Based Statistical Area (CBSA). CBSAs are non-overlapping geographic areas defined by the U.S. Office of Management and Budget that exhibit market-related behaviors. CBSAs are socially and economically interdependent areas that are geographically circumscribed by commuting times to the core geographic area. For the few providers that were not located within a CBSA, we assigned them to the largest CBSA located within their Hospital Referral Region (HRR), which are regional health care markets for tertiary hospital care.

Based on our market definition, we excluded markets from our comparison group if they were very different from the markets with BPCI participants active during the first four quarters. To this end, we first characterized BPCI markets by urban location and population size, and identified which market types did not have any BPCI participants. For example, Model 2 acute care hospital (ACH) providers participating in spinal surgery episodes were located exclusively in urban counties. Thus, we excluded ACHs from the comparison group located in rural markets.

Additional exclusions were imposed to control for differences in regulatory environments that may relate to differences in outcomes. For instance, for Model 2 we excluded all markets in the

---

[20] We had an insufficient sample of Model 3 IRF and LTCH EI types to conduct propensity or Mahalanobis matching. Therefore, Step 2 was not conducted for these EI types.

state of Maryland, because Maryland hospitals are not paid under Medicare's Inpatient Prospective Payment System (IPPS).

We excluded non-BPCI providers in markets where BPCI participants have over 50% of the discharges in the 48 BPCI clinical episodes. This was to avoid including providers that may be exposed to spillover effects of BPCI. The presence of a BPCI participant in a market may cause changes in utilization or referral patterns for other beneficiaries in the market. This spillover effect may confound interpretation of results because non-BPCI beneficiaries may receive some care from BPCI participants, comparison providers may adopt practices similar to BPCI participants, or BPCI may affect referral patterns in the market. Exhibit 13 includes all market-level characteristics that were considered when identifying eligible comparison group providers by Model and by provider type.

Provider exclusions

We also excluded providers with characteristics that were not exhibited by BPCI providers. For example, based on Model 2 and 4 ACH participant characteristics, only comparison providers meeting minimum values for number of discharges in 2011 were included in the propensity score or Mahalanobis model. The minimum values were calculated by rounding down the minimum value for the ACH BPCI participants in each clinical group. Exhibit 13 includes all provider characteristics that were considered when identifying eligible comparison group providers by Model and by provider type.

Exhibit 13: Market-level and provider characteristics used in determining non-BPCI provider eligibility for comparison group, by model and EI type

Model	Episode Initiator Type	Market-level characteristics	Provider characteristics
Models 2 & 4	ACH	▪ Urban vs. rural CBSA ▪ Exclude Maryland ▪ Market share of BPCI providers (>50%)	▪ Ownership (government; non-profit; for-profit) ▪ Exclude if in same organization as a BPCI participant (same "owner") ▪ Exclude if in Phase 1 with a parent in Phase 2 ▪ Exclude future BPCI participants
Model 3	SNF and HHA	▪ Urban vs. rural CBSA ▪ Market share of BPCI providers (>50%)	▪ Ownership (government; non-profit; for-profit) ▪ Exclude if in same organization as a BPCI participant (same "owner") ▪ Exclude if in Phase 1 with a parent in Phase 2 ▪ Exclude future BPCI participants
Model 3	IRF and LTCH	▪ State (must be BPCI provider in the state)	▪ Exclude if in same organization as a BPCI participant (same "owner") ▪ Exclude if in Phase 1 with a parent in Phase 2 ▪ Exclude future BPCI participants

Step 2: Use propensity and Mahalanobis distance to identify close matches

In addition to the market and provider attributes described above, BPCI participants differed from non-participants in less obvious ways. Overall, BPCI participants were located in markets with greater competition, but some BPCI participants were located in markets with little competition.[21] Model 2 BPCI participants were located in all regions (Midwest, Northeast, South, and West), but relative to all ACHs, there were fewer participants in the South and more in the Northeast. Relative to non-participants, BPCI participants were more likely to be located in markets with more providers per capita, including specialists and SNFs, and to have higher Medicare Part A payments per admission prior to joining BPCI, but there was significant variation across BPCI participants.

Because of these differences, we used propensity score and Mahalanobis methods to identify adjustments to the comparison groups to help ensure a better match. *Propensity score* is defined as the probability of receiving the treatment (in this case, participating in BPCI), conditional on a set of characteristics. This probability was estimated using a logistic model that included key factors of interest defined in conjunction with CMS. The factors included market characteristics (i.e., population size, SNF beds, primary care physicians), provider characteristics (i.e., part of a chain, size), and historical Part A Medicare payments and patient PAC use. Using the coefficients from the logistic regression model, we constructed a propensity score as the predicted probability of participating in BPCI. We used the calculated propensity score to match BPCI providers to their 15 nearest neighbors with replacement. We had sufficient sample size to use the propensity score model to identify the 15 nearest neighbors for Model 2 BPCI providers participating in orthopedic surgery episodes. For the remaining Model/EI type/clinical episode groups, we did not have sufficient sample size to estimate a propensity score model to calculate a propensity score and used Mahalanobis distance matching instead. We first calculated the Mahalanobis distance for each BPCI participant and non-BPCI participant pair in our sample. The Mahalanobis distance between two set of covariates X_i and X_j, is defined as $d_{ij} = (X_i - X_j)'S^{-1}(X_i - X_j)$ where S^{-1} is the inverse of the correlation matrix of the variables included in vector X. We then matched each BPCI provider to the 15 participants in the non-BPCI group with the lowest Mahalanobis distances.

To select an optimal matching model, we conducted a series of diagnostic tests for each candidate model to assess whether BPCI providers are comparable on all observed covariates to the matched sample of comparison providers. Some diagnostics included conducting t-test for statistical significance of difference in means in all covariates, and the standardized differences between the BPCI participants and the comparison groups in all characteristics. The model selection required multiple iterations with the goal of reducing the standardized differences below 0.2 for the maximum number of key market and provider characteristics. **Appendix J** shows the standardized differences of each covariate included in the model between BPCI providers and matched comparison providers for each Model/EI type/clinical episode group. As shown in Appendix J, our ability to construct comparison groups varied across Model/EI type/clinical episode groups. Overall, we were able to reduce standardized differences in a larger

---

[21] We measured the competition of a market using the Herfindahl index. The Herfindahl index is defined as the sum of the square of the market shares (i.e., market penetration) of all providers (BPCI and non-BPCI) of a particular type (ACH, SNF, HHA, etc.). The Herfindahl Index values can range from 0 to 1, where values closer to zero signify a higher degree of competition among providers and values closer to 1 signify less competition (i.e., one or a few providers dominate the market).

number of characteristics in Model/EI type/clinical episode groups with the largest number of participants and with the most homogenous clinical groups. Comparison hospitals for several of the Model 2/clinical episode combinations, those with the most disaggregated clinical groups and largest number of participants per group, show small standardized differences on the vast majority of the market and provider characteristics and baseline outcomes, such as readmission rates and Medicare payments. However, comparison providers for most of the Model 3 and Model 4, with the exception of Model 3/HHA/non-surgical respiratory group, were matched to BPCI participants on a very small set of variables due to the small number of participants.

Step 3: Construct episodes for BPCI and matched comparison providers

Once we selected the final BPCI and comparison providers for a given quarter, we use the BPCI rules to identify all episodes of care that began between October 2010 and September 2014 and that would have been assigned to comparison facilities if they were participating in BPCI.

Step 4: Select random sample of comparison group episodes

Among all episodes identified in the previous step, we drew a sample to match the distribution of MS-DRGs and dates of episode start in the BPCI population. To this end, each BPCI episode was randomly matched to one episode from a group of episodes in the same quarter and either the same MS-DRG (for ACH's) or MS-DRG group (for SNFs and HHAs), originating from the comparison providers that were matched to the provider generating the BPCI episode. This episode was then excluded from the pool of episodes eligible for future matching any additional BPCI participant episodes.

3.   Beneficiary survey sample

This annual report includes the results for two survey waves that were conducted between June 2014 and November 2014. In this section, we describe the creation of the samples used for both Wave 1 (May and June 2014) and Wave 2 (October and November 2014) BPCI and comparison survey samples.

a.   Sampling frame

For Wave 1 we constructed the Model 2 sample using fee-for-service (FFS) claims from June 2014, based on the hospital discharge date. For Wave 2 we constructed the sample using Medicare claims from two "rolling" one-month samples, which received their first surveys a month apart.[22] For the first rolling month of Wave 2, claims for October 2014 were pulled in early November 2014 and surveys were mailed in the first week of January 2015. For the second rolling month of Wave 2, claims for November 2014 were pulled in early December 2014 and surveys were mailed the first week of February 2015. This rapid sampling process was deliberately used to reduce

---

[22] While one-month of claims was adequate to fill all Model 2 strata in Wave 1, it took two months of claims to fill all Model 2 strata in Wave 2. This is because the Wave 2 strata were more specific with regards to patient diagnosis, which meant there were fewer eligible episodes in each of the two months.

recall bias. It does, however, limit the sample to patients whose claims were filed quickly, within one month of discharge.[23]

For both Wave 1 and Wave 2, we constructed the sample for Model 3 strata using two months of Medicare claims. These were drawn in a single data pull, which encompassed claims for PAC admissions in the prior two months. We did not use the strategy of two rolling months because PAC claims generally take longer to process. If we used a rolling strategy, our two combined one-month samples would be smaller than one combined two-month sample, and we would risk falling short of the sample size necessary for acceptable statistical precision. Exhibit 14 summarizes the periods used to create the sample frame of the BPCI beneficiary survey Waves 1 and 2.

Exhibit 14: Episode period for beneficiary survey sample frame,
Models 2 and 3, Waves 1 and 2

Wave	Model 2	Model 3
Wave 1: Summer 2014	One one-month period: June 2014 hospital discharges	One two-month period: May & June 2014 admissions to PAC
Wave 2: Fall 2014	Two "rolling" one-month periods: October & November 2014 hospital discharges	One two-month period: October & November 2014 admissions to PAC

b. Sample construction

The first step to construct the survey sample was to exclude some providers from the sample frame to maximize efficiency and validity of the survey analysis. We used the provider characteristics available from the Medicare claims and administrative data to compare the BPCI providers with a similar set of providers that were not participating in BPCI. The provider characteristics included provider type (acute care hospital, SNF, HHA), provider size (small vs. large), academic affiliation, ownership type (for-profit, non-profit and government/other), census region, and urban/rural location. The combination of all provider characteristics yielded a maximum of 96 "provider strata" for each model. Some of these provider strata were empty (no providers) in the BPCI or comparison group. We dropped all strata that were represented in one group but not the other because it was not possible to create matched comparison groups for BPCI providers in these strata.

The second step was to create cells within each of the sampling strata and then match beneficiaries within each cell by provider and patient characteristics. We began by ordering the factors that could be used for defining the cells in order of most to least importance in terms of their influence on patient outcomes. The order was as follows: provider type, BPCI episode groups, patient age groups, provider size, academic affiliation, ownership type, urban/rural indicator, census region, and beneficiary gender. When defining the cells, we aimed to strike a balance in the survey sampling design that: 1) matched on factors that we thought would most affect beneficiary outcomes; and 2) had a sufficient number of episodes in each stratification cell to support valid comparisons. At a minimum, we required that every cell had at least 10 discharges for BPCI providers and 10 for comparison providers.

---

[23] Although claims submitted within one month may not represent the entire Medicare population within a stratum due to provider delays in submitting claims, this issue should affect BPCI and comparison samples equally, and not bias our estimates.

C. Outcome Measures

In this section we define the nine clinical episode groups we created from the 48 clinical episodes, the various measurement periods during which we define the outcome measures, and summarize the outcome measures presented in the results section.

1. Clinical episode aggregation

During the first year of the initiative, there were not sufficient sample sizes to report outcomes by model and by clinical episode. To accommodate small sample sizes, we consulted with clinicians at Telligen, and CMS, to group clinically similar episodes. Our goal was to aggregate as little as possible while ensuring clinically meaningful subgroups of clinical episodes with sufficient sample sizes. In conversations with CMS and Telligen, we decided nine clinical episode groups would be the most appropriate level at which to stratify the results by Model, if sample sizes were sufficient. Exhibit 15 provides a description of each clinical episode group and one or two representative clinical episodes included in the clinical episode group. See **Appendix K** for a detailed table depicting how we combined the 48 clinical episodes into nine clinical groups.

Exhibit 15: Nine Clinical Episode Groups and Representative Clinical Episodes

Clinical Episode Groupings	Representative Clinical Episodes included in the Clinical Episode Group
Non-surgical and surgical gastrointestinal	Esophagitis, gastroenteritis and other digestive disorders
Non-surgical cardiovascular	Congestive heart failure
Non-surgical neurovascular	Stroke
Non-surgical orthopedic	Fractures of the femur and hip or pelvis
Non-surgical other medical	Sepsis
Non-surgical respiratory	Chronic obstructive pulmonary disease, bronchitis, asthma
Cardiovascular surgery	Coronary artery bypass graft
Orthopedic surgery	Major joint replacement of the lower extremity
Spinal surgery	Cervical spinal fusion

2. Measurement periods

For this evaluation, we defined two sets of *measurement periods* for which we calculated the outcomes of interest: the *bundle timeline* and the *patient timeline*. The bundle timeline measurement periods vary by model and by episode length. In contrast, the patient timeline measurement periods are consistent across models and episode lengths. This allows us to compare outcomes regardless of the episode lengths and models. Every outcome was calculated for one or more defined *measurement periods.* For example, for Models 2 and 4, all-cause, unplanned readmission rates were calculated for three *patient timeline* measurement periods: within 30 days, within 60 days, and within 90 days of hospital discharge. These measurement periods are labeled *post-discharge 30, post-discharge 60, and post-discharge 90.* Exhibits 16 and 17 describe the bundle timeline measurement periods and the patient timeline measurement period.

Episodes were dropped from measure denominators on a case-by-case basis in situations where there was not enough claims run-out to cover the measurement period. Specifically, if the end of our current observational period (December 31, 2014) occurred within the *measurement period* for

the given episode, we dropped the episode from the denominator. For example, if a Model 2 episode began on September 23, 2014 and had a post discharge period beginning October 6, 2014, we dropped the episode from any 90-day post discharge measures, since the 90-day post discharge period for this episode extends beyond December 31, 2014. As a result of these exclusions, the outcomes measured during the 90-day post discharge *patient timeline* period (or episode start plus 90 days) have smaller denominators than outcomes measured during the 30-day post discharge *patient timeline* period (or episode start plus 30 days).

Exhibit 16: Definition of Measurement Periods Relative to the Bundle across Models

		Bundle Dates		Within Bundle Services					
Definition of Measurement Periods Relative to the Bundle									
Model	**Pre-bundle 30**	**Start date**	**End date**	**Acute**	**Post-Discharge**	**Episode Start +30, +60, and +90**	**Post PAC 30**	**Post-bundle (PB) 30**	**Post-bundle (PB) 60**
Model 2	Anchor IP stay admission date minus 30 days	Anchor IP stay admission date	Anchor IP stay discharge date plus bundle length	Anchor IP stay from IP admission date to IP discharge date[b]	From IP discharge date to bundle end date	Anchor IP stay admission date plus 30 (60, or 90) days	NA	30 days after the end of the bundle	31 to 60 days after the end of the bundle
Model 3	EI PAC admission date minus 30 days	EI PAC admission date	EI PAC admission date plus bundle length	N/A	From EI PAC discharge date to bundle end date	EI PAC admission date plus 30 (60, or 90) days	EI PAC discharge date plus 30 days	30 days after the end of the bundle	31 to 60 days after the end of the bundle
Model 4	Anchor IP stay admission date minus 30 days	Anchor IP stay admission date	IP stay discharge date (anchor IP stay if no readmission occurs OR qualifying readmission)[a]	Anchor IP stay from IP admission date to IP discharge date[b]	Duration of qualifying readmissions started within the 30-day readmission window	Anchor IP stay admission date plus 30 (60, or 90) days	NA	30 days after anchor IP discharge date excluding days related to qualifying readmissions	31 to 60 days after anchor IP discharge date excluding days related to qualifying readmissions

Notes:

[a] If a qualifying readmission occurs within 30 days after anchor admission discharge date, the period between anchor hospital discharge and hospital readmission date belongs to the post-bundle period.

[b] For BPCI beneficiaries who were transferred from an anchor hospital to another hospital, the acute care period ends at the discharge date from the transfer hospital.

Exhibit 17: Definition of Measurement Periods Relative to the Patient Timeline across Models and episode lengths

Model	Pre-Admission	Anchor IP	30-day Post-Discharge Period (PDP)	60-day PDP	90-day PDP
Model 2	30 days prior to anchor hospital stay	Anchor IP stay from IP admission date to IP discharge date	PDP from anchor IP discharge date to 30 days	PDP from anchor IP discharge date to 60 days	PDP from anchor IP discharge date to 90 days
Model 3					
Model 4					

3. Outcome definitions

In this section we present the outcome measures that were constructed and analyzed to evaluate the impact of BPCI during the first year of the initiative. Exhibit 18 summarizes the key outcome measures, organized by domain, that are presented in the results section of the report. **Appendix L** provides detailed definitions of each outcome measure.

Exhibit 18: Quantitative Outcome Measures used to Evaluate the Impact of BPCI organized by Domain and Data Source

	Domain/Quantitative Outcomes	Medicare Claims	MDS assessments	OASIS assessments	IRF-PAI assessments	Beneficiary Surveys
Payment	Total Medicare standardized allowed payment for inpatient stay plus 90 days post-discharge	X				
	Total Medicare standardized allowed payment included in the bundle definition	X				
	Total Medicare standardized allowed payment not included in the bundle definition	X				
	Medicare standardized allowed payment, 30-day pre-bundle period	X				
	Medicare standardized allowed payment, 30 and 60 post-bundle period	X				
	Total Medicare Part A standardized allowed payment (by various settings)	X				
Utilization	Acute inpatient length of stay	X				
	Number of days in PAC setting (total and by setting)	X				
	Number of home health visits	X				
	First PAC setting following inpatient discharge	X				
	Discharged to institution relative to discharged home with home health	X				
Quality	Unplanned readmission rate following inpatient hospital discharge	X				
	Emergency department use without hospitalization	X				
	All-cause mortality	X				

	Domain/Quantitative Outcomes	Medicare Claims	MDS assessments	OASIS assessments	IRF-PAI assessments	Beneficiary Surveys
Functional status	SNF PAC setting:[b,c]					
	% of SNF patients who improve status or remain completely independent in long-form ADL function (a measure of overall function))		X			
	% of SNF patients who improve status or remain completely independent in early-loss ADL function (a measure of self-care function)		X			
	% of SNF patients who improve status or remain completely independent in mid-loss ADL function (a measure of mobility)		X			
	HHA PAC setting:[b,d]					
	% of HHA patients who improve status or remain completely independent in bathing; upper body dressing; lower body dressing; ambulation/locomotion; bed transferring			X		
	IRF PAC setting:[b]					
	Average Changes in Self-Care Score				X	
	Average Changes in Mobility Score				X	
	% Patients with improvement in bathing, dressing, using the toilet, or eating					X
	% Patients with improvement in walking without rest					X
	% of patients with improvement in use of mobility device (i.e., less frequent)					X
	% of patients with improvement in using stairs					X
	% of patients with improvement in planning regular tasks					X
	Improvement in physical/emotional problems limiting social activities (i.e., less frequent)					X
	Improvement in pain limiting regular activities (i.e., less frequent)					X
Patient Experience	% of patients who have limited normal activities because of pain					X
	How often did you, your family, or your caregiver get conflicting advice from medical staff about your treatment					X
	How often were the services you got appropriate for the level of care you needed					X
	% of patients who thought they were discharged at the right time					X
	% of patients who had a good understanding of how to take care of themselves before they prepared to leave the hospital					X
	% of patients who thought medical staff clearly explained how to take your medications prior to leaving hospital					X
	% of patients who thought medical staff clearly explained what follow-up appointments or treatments would be needed					X
	% of patients who thought the medical staff took the patient's preferences into account in deciding what health care services they should have after they left the hospital					X
	% of patients who since having left the hospital thought that they and their caregivers have been able to manage their health needs					X

	Domain/Quantitative Outcomes	Medicare Claims	MDS assessments	OASIS assessments	IRF-PAI assessments	Beneficiary Surveys
Market Dynamics	PAC referral network concentration index (Models 2/4)	X				
	Hospital referral network concentration index (Model 3)	X				
	Number of PAC providers receiving referrals, per Model EI (Models 2/4)	X				
	Number of hospitals referring patients, per Model EI (Model 3)	X				
	Highest percent of referrals to one PAC provider (Model 2)	X				
	Highest percent of referrals from one hospital (Model 3)	X				
	Percent of market share	X				

Notes: Descriptions of qualitative data are located in the Data Sources section above.

[a] These amounts combine the Medicare payments with the patient coinsurance and copayment amounts and then adjust for Medicare payment policies to ensure that any differences across time and providers reflect real differences in resource use rather than Medicare payment policies (e.g. teaching payments or differential payment updates).

[b] For BPCI Models 2 and 4, the eligible sample for the functional status measures is based on the first PAC setting (SNF, HH, or IRF) to which a patient was discharged after the inpatient stay that triggered an episode of care (the "anchor hospitalization"). For BPCI Model 3, the approach focuses on the patients' first encounter with a BPCI-participating PAC provider after the anchor hospitalization. We only included the first PAC stays with a valid beginning assessment within 30 days after discharge from the anchor hospitalization (i.e., anchor discharge) and a valid final assessment within 120 days after the anchor discharge.

[c] For SNF, we used the 5-day assessment as the beginning assessment and the discharge assessment or the latest available assessment within 120 days after anchor discharge as the ending assessment.

[d] For HHA, we used the start of care assessment as the beginning assessment and the discharge assessment or re-certification assessment within 120 days after anchor discharge as the final assessment.

D. Analytical Methods

To answer the research questions listed above, we used a mixed method approach using both quantitative and qualitative analytical methods.   These methods varied depending on the research question and the data used to calculate measures as shown in **Exhibit 19**.   These methods include descriptive analysis, DiD, cross-section comparisons between beneficiaries, BPCI, and comparison survey respondents, before-after comparisons across BPCI participants, analysis of market dynamics, and qualitative analysis. This section outlines our approach for each of these analyses.

Exhibit 19: Crosswalk of Research Questions to Analytical Methods

Research Question	Descriptive Analysis	DiD	Cross-section comparisons between BPCI and non-BPCI survey samples	Before-After Comparisons Across BPCI participants	Market Dynamic Analysis	Qualitative Analysis
A. What are the characteristics of the program and participants at baseline and how have they changed during the course of the initiative?	X					X
B. What is the impact of the BPCI initiative on the costs of episodes, the Medicare program, and the quality of care for Medicare beneficiaries?						
Impact on payment and utilization		X				
Impact on quality of care		X	X			X
Impact on provider referral an market share					X	
C. What program, provider, beneficiary, and environmental factors contributed to the various results of the BPCI initiative?	X			X		X

1. Descriptive analysis

To summarize characteristics of the program and participants at the baseline and during the course of the initiative (Research Question A), we run a series of descriptive analyses on measures drawn from BPCI Implementation Packages, Quarterly Reported Awardee data, POS files, and the AHRF. Cross-tabulations are created to provide insight into potential associations between provider/Awardee characteristics and cost and quality outcomes of interest.

2. Difference-in-differences analysis

The DiD approach quantifies the impact of BPCI by comparing changes in claims and assessment-based outcomes for the BPCI population with changes in outcomes for the comparison population, between the baseline and intervention periods. This approach eliminates biases from time invariant differences between the BPCI and comparison populations, and controls for trends in the BPCI population. To mitigate selection bias, the DiD model incorporates outcomes from three periods prior to BPCI implementation (before, baseline, and Phase 1) as well as the intervention period to control for time invariant differences in the mean outcomes between the

two groups.[24] Because BPCI participants started implementing changes during Phase 1 in preparation for the risk-bearing phase (the intervention), the Phase 1 period was excluded from the baseline, as well as from the intervention period. Thus, the DiD compares changes in outcomes from the baseline period to the intervention period.

- The DiD baseline period was from October 2011 through September 2012.
- The transition period (Phase 1) was from October 2012 through September 2013.
- The BPCI to date intervention period was from October 2013 through September 2014. [25]

To illustrate the calculation of the DiD in a regression framework, consider the linear model listed below:

$$Y_{i,k,t} = \alpha + \beta_1 BPCI_{i,k,t} + \beta_2 T_t + \delta BPCI_{i,t} \cdot T_t + \boldsymbol{X_{i,k,t}}'\beta + u_{i,k,t}$$

Where $Y_{i,k,t}$ is the outcome of interest for individual *i* with provider *k* in quarter *t*, $BPCI_{i,t}$ is an indicator variable taking the value of 1 if individual *i* was treated by a BPCI provider, T_t indicates the period (i.e., before, baseline, transition, or intervention), and $X_{i,k,t}$ are beneficiary demographics and clinical characteristics observed before hospitalization, and provider characteristics. The vector β is a vector of regression coefficients that captures the impact of risk factors $\boldsymbol{X_{i,k,t}}$ in the outcome of interest. The regression coefficient β_1 captures any inherent, time invariant differences between the control and the treatment groups, while β_2 provides an estimate of the potential time trends in the outcome of interest over the period before and after the intervention that is common to both the control and treatment groups, while u_{ikt} represents a random error term. In this linear example, the DiD estimate is simply coefficient δ, which determines the differential in outcome Y experienced by beneficiaries receiving services from BPCI providers during the intervention period relative to beneficiaries receiving services from providers in the comparison group.

We used multivariate regression models to control for differences in beneficiary demographics and clinical characteristics observed before the hospitalization, along with prior care use, and provider characteristics that might be related to the outcome. We used a common set of variables in all of our models for simplicity and ease of data collection and analysis. For example, all measures were risk-adjusted for service mix using MS-DRG information from the episode triggering inpatient stay (Model 2 and Model 4) or qualifying inpatient stay (Model 3). Demographic factors included in all models are age brackets, gender, age and gender interactions, Medicaid eligibility status, and disability status. To control for prior health conditions, we used Hierarchical Chronic Conditions (HCC) indicators[26], which could be used individually or

---

[24] While the DiD model controls for unobserved heterogeneity that is fixed over time, there is no guarantee that this unobserved heterogeneity is, in fact, fixed. It could be the case, for example, that providers with improving outcomes are relatively more likely to sign up for the program inducing a spurious positive correlation between BPCI participation and outcomes. Future developments of the comparison group of providers will use information on historical trends in outcomes as matching variables.

[25] Assessment-based outcomes including functional status outcomes and post-bundle payment outcomes are reported with one quarter delay. The DiD results for these outcomes use October 2013 to June 2014 as an intervention period. All other periods remain the same.

[26] The hierarchical condition categories (CMS-HCC) model is a prospective risk-adjustment model used by CMS to adjust Medicare Part C capitation payments for beneficiary health spending risk. The model adjusts for demographic and clinical characteristics. The clinical component of the model uses diagnoses from qualifying services grouped

aggregated. To further control for case-mix differences, we include measures of prior care use in the following settings: hospital, LTCH, SNF, institutional stay, IRF, hospice, HHA, psychiatric facility, and emergency department. In addition, to account for regional differences we either used a set of provider characteristics including region, bed count, and for-profit status, or used state dummies.

While the same demographic and enrollment status indicators are included for all outcomes, we considered alternative specifications to control for service mix, clinical factors, prior care use, and regional characteristics. These are listed in Exhibit 20. To assess different specifications, we split the sample into a model development and a validation sample, and estimated each model using data from the model development sample. We then evaluate models in terms of their goodness of fit (Akaike Information Criterion (AIC) and Bayesian Information criterion (BIC) criteria, R-square, t-tests on differences in conditional expectations by subgroup) in the model development sample and their predictive performance in the validation sample. Once risk-adjustment variables were selected, we added quarterly indicators that interacted with an indicator on BPCI participation to our models. Risk adjustment model specifications for each outcome and model are listed in **Appendix M.**

Exhibit 20: Predictive Risk Factors Used to Risk-Adjust Outcomes

Domain	Variables
Service Mix	Alternative specifications ▪ Anchor MS-DRG ▪ MS-DRG group: anchor MS-DRG grouped with and without complications together ▪ 48 clinical episodes ▪ Clinical groups (see Section II.C.1)
Patient Demographics and Enrollment	▪ Age (under 65, 65-79, 80+) ▪ Gender ▪ Medicaid status ▪ Disability status
Clinical Factors	Alternative specifications ▪ HCC indicators from qualifying services and diagnoses from claims and data for months preceding the anchor admission or qualifying stay ▪ HCC aggregated to 45 risk variable groups (RV-HCC) according to NQF measure 1789 (**Appendix N** shows a crosswalk from HCC groups to RV-HCC.) ▪ HCC index, HHC indicators weighted by their relative weight in the CMS-HCC model

---

into several HCC indicators. Pope, Gregory C.; Kautter, John; Ellis, Randall P.; Ash, Arlene S.; Ayanian, John Z.; Iezzoni, Lisa I.; Ingber, Melvin J.; Levy, Jesse M.; and Robst, John, "Risk adjustment of Medicare capitation payments using the CMS-HCC model" (2004). *Quantitative Health Sciences Publications and Presentations.* Paper 723.

Domain	Variables
Utilization measures preceding the start of the anchor stay/qualifying inpatient stay	Alternative specifications ▪ Binary indicators for utilization of ED, inpatient, SNF, IRF, HHA services in the **six months** preceding the start of the episode, and ever in a nursing facility (NF)/SNF in the six months preceding the start of the episode ▪ Number of days of ED, inpatient, SNF, IRF, HHA service use in the **one month** preceding the start of the episode, and ever in a NF/SNF in the six months preceding the start of the episode ▪ Number of days of ED, inpatient, SNF, IRF, HHA service use in the **six months** preceding the start of the episode, and ever in a NF/SNF in the six months preceding the start of the episode
Market Factors	▪ Managed care penetration ▪ Median household income in the market ▪ State indicators ▪ Census region indicators
Provider Characteristics	▪ Size ▪ Ownership status

We used a variety of empirical specifications including ordinary least squares (OLS) and logistic regressions, duration and two-part models. Models were estimated depending on the type and characteristics of the outcome measure. For example, logistic models were estimated for the discrete quality outcomes, (e.g. mortality, readmission rate, and ED use). A Cox proportional hazard model was used to estimate inpatient stay. Ordinary least square models (OLS) were estimated for the continuous total number of days measures (e.g. number of days during the anchor hospitalization or number of SNF days) as well as some of the payment models including total payments that were covered by the bundle where all individuals by default had positive expenditures. Two part models were favored for payment outcomes where more than 5% of individuals had zero payments for the particular outcome. These payment outcomes included the individual part A and part B payments that were affected by zero-mass and skewness.

Estimates from the multivariate regression models were used to construct model-predicted outcomes under two scenarios (baseline and intervention) for both BPCI-participating and comparison hospitals. To control for changes in service and case-mix over time as well as differences between BPCI and non-BPCI beneficiaries, we used the same reference population of beneficiaries to calculate quarterly predicted outcomes for BPCI providers and providers in the comparison group. The reference population used in this report is all beneficiaries treated by BPCI providers during the intervention period. We tested for equality of trends in key outcomes between the BPCI participants and comparisons and found that the trends were the same. Therefore, we conclude that this main DiD assumption was met, so our estimates are unlikely to biased by pre-BPCI trends in the key outcomes considered.

The DiD estimate was then calculated by first taking the difference between the two scenarios for both BPCI-participating and comparison hospitals and thereafter taking the difference between BPCI-participating and comparison hospitals. Taking the average difference in such differentials across all BPCI beneficiaries yields the Effect of the Treatment on the Treated (ETT) analog of the

DiD. The ETT is the average gain from treatment for those who actually were treated. Standard errors of ETT estimation were computed using the Delta method.[27]

We report DiD estimates for each given outcome, if the sample exceeds 30 BPCI episodes in the first quarter of the intervention period for outcomes evaluated using duration, logistic, and OLS models. In contrast, we used a minimum of 100 BPCI episodes in the first quarter of the intervention period to report DiD estimates for outcomes using two-part models. For all clinical episode groups we have most DiD estimates reported. Some outcomes including number of IRF days and number of LTCH days and payment outcomes that are stratified by bundle length and PAC status inherently suffer from small sample sizes and consequently, DiD estimates for these particular outcomes were largely not reported.

3. *Cross-section comparisons between BPCI and comparison survey respondents*

a. *Analysis of functional questions*

For the functional measures in Waves 1 and 2[28], we calculated the risk-adjusted rate of improvement and decline in functional status for both the BPCI respondents and the comparison respondents and estimated the treatment effect (i.e. difference in rates of change between the two groups) and its 95% confidence interval. Analyses in Wave 1 were conducted for each sampling stratum within each model (MCC and non-MCC), as well as the pooled model level. Analyses in Wave 2 were conducted for each sampling stratum (major joint replacement of the lower extremity, nonsurgical cardiovascular, nonsurgical respiratory and all Model 3 episodes).

We asked respondents to recall their functional status before their treatment episode, and also to report their functional status on the day they filled out the survey. For each functional status measure, we created binary indicators for both improvement and decline to measure change before and after the treatment episode. The improvement indicator takes a value of 1 if a patient moved to a better functional group after the hospitalization (e.g. from "complete help needed" before to "no help needed" after the hospitalization), or if the respondent recalled having the best functional status prior to hospitalization and remained in the best status at the time of survey response (i.e. "no help needed" both before and after the hospitalization). The indicator is assigned a value of 0 otherwise. The decline indicator takes a value of 1 if the patient moved to a worse functional group

---

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

[28] The seven functional measures in both waves were: (1) bathing/dressing/toileting/eating; (2) planning regular tasks; (3) moving using a mobility device; (4) walking without resting; (5) going up or down stairs; (6) the frequency with which physical or emotional health interferes with regular social activities , and (7) the frequency with which pain interferes with normal activities. For the first five functional measures, patients were asked to recall their functional status before the anchor hospitalization and also to report their current status at the time of the survey in both waves. In Wave 1, the survey asked about measures (6) and (7) only at the time the survey was being filled out and did not ask respondents to recall their status prior to hospitalization. For Wave 2 we asked about prior and current status for all seven measures.

after hospitalization, or if the patient recalled having the worst functional status prior to hospitalization and remained in the worst status at the time of filling out the survey.

b. Analysis of other outcomes

For all other questions in the survey (health care experience questions), we created a dichotomous variable for each question by collapsing the response set. We picked the cut-off for each question based on logical dichotomies in the response options (e.g., agree/disagree) and the distribution of responses. The results tables in **Appendix O** provide details regarding the collapse of response set for each question. We then used logistic regression to calculate the rates of the dichotomous indicators for the BPCI and comparison respondents, and used these to estimate the treatment effect and its 95% confidence interval.

All analyses (functional measures and other measures) in Wave 1 were conducted for each sampling stratum within each model (MCC and non-MCC), as well as the pooled model level. All analyses in Wave 2 were conducted for each sampling stratum (major joint replacement of the lower extremity, nonsurgical cardiovascular, nonsurgical respiratory and all Model 3 episodes) but not at the pooled level. This is because, in Wave 2, we only sampled Model 2 beneficiaries with select clinical episode groups; so analysis of all Model 2 strata combined does not reflect the overall BPCI program effect on all Model 2 beneficiaries. In Wave 1, however, we sampled Model 2 beneficiaries by their MCC status and did not restrict the sample to beneficiaries with certain clinical episode groups.

c. Weighting

We applied both sampling weights and nonresponse weights when conducting survey data analysis.[29] The sampling weight is the inverse of the selection probability within each of the 148 or 172 sampling strata for Waves 1 and 2, respectively. The nonresponse weight was calculated for all survey respondents (complete and partial responses) and reflects the inverse of the probability of response among eligible members of the sample (with deceased respondents removed) within each of the 148 (172) sampling strata. The final nonresponse-adjusted weight was calculated as the product of the sampling weight and the nonresponse weight.

Under perfect conditions (i.e. no decedents, no item nonresponse), use of the nonresponse-adjusted weight would balance the BPCI and the matched comparison sample on the variables used to define the sampling strata. However, differential ineligibility (i.e., death rates) and nonresponse on any particular survey question can create imbalance, requiring us to control for some of the variables used to define the cells (e.g., age).

d. Risk adjustment

Although the BPCI and comparison survey samples were matched by important patient and provider characteristics, risk adjustment is still necessary for some outcome measures to control for other important risk factors, such as baseline functional status, to ensure comparability, as much as possible, between the two groups. Without adequate risk adjustment, it would not be

---

[29] For both BPCI and comparison respondents, the sampling weights sum to the population size of the BPCI group.

possible to determine if different outcomes were due to the impact of BPCI or to differences in the survey samples.

We performed regression-based risk adjustment for survey questions related to patient physical and mental health. These questions included the functional improvement/decline measures and the additional health status questions discussed in Section II.D.3.b above. The risk factors we controlled for included beneficiary age, gender, Medicare/Medicaid dual eligibility, clinical episode groups, Hierarchical Condition Categories (HCC) score, the number of SNF and inpatient hospital days in the 90 days prior to the anchor hospitalization, whether the beneficiary was admitted to the anchor hospital from the community setting or from a nursing home, and most importantly, the beneficiary's recalled functional status prior to the treatment episode. [30]

For analyses of Model 3 episodes, we controlled for two additional risk factors. The first was whether the episode-initiating provider was an HHA or a SNF. The second was the number of days that elapsed between discharge from the anchor hospital stay and admission to SNF or HHA care. This gap between hospital discharge and PAC admission may indicate incomplete care, and may also indicate the emergence of new, unrelated medical conditions.

Note that none of the questions related to health care experiences were risk adjusted. This is because all beneficiaries should receive the same high level of care in the hospital and afterward, regardless of their demographics or clinical risk factors. For all multivariate analyses, we estimated robust standard errors.[31]

4. Before-after comparisons across BPCI-participating providers

To investigate market, provider, and program factors that may lead to differences in observed changes in outcomes from baseline to intervention period across BPCI providers, we compared changes in risk-adjusted outcomes among providers in the same Model and clinical group. The study population includes BPCI providers during Q4 2013 through Q2 2014 that had at least 25 episodes in both the baseline and intervention periods. We only report results for Models and clinical groups for which at least 20 BPCI providers meet the inclusion criteria and for which we found statistically significant results between BPCI and the comparison group under research

---

[30] Because Wave 1 strata were defined at the MCC/non-MCC level we used the following nine aggregate clinical episode groups as controls: orthopedic surgery excluding spine, non-surgical: other medical, non-surgical: neurovascular, non-surgical: cardiovascular, non-surgical: respiratory, non-surgical and surgical: GI, cardiovascular surgery, non-surgical: orthopedic, and spinal surgery. In Wave 2, Model 2 strata were already defined at BPCI episode group or aggregate episode group level. For Model 2 MJRLE episodes, we did not have any episode group controls because the stratum was defined at the BPCI episode group level. For Model 2 "non-surgical cardiovascular" episodes, we controlled for seven BPCI episode groups: congestive heart failure, cardiac arrhythmia, acute myocardial infarction, chest pain, medical peripheral vascular disorder, syncope/collapse, and atherosclerosis. For Model 2 "non-surgical respiratory" episodes, we controlled for three BPCI episode groups: chronic obstructive pulmonary disease/bronchitis/asthma, other respiratory, and simple pneumonia and other respiratory infection. Model 3 episodes continued to be controlled for using the nine aggregate clinical episode groups used in Wave 1.

[31] We explored the possibility of clustering at the provider level. However, the median provider only contributes one observation to the sample, suggesting there is no need to cluster. Additional testing did not reveal any meaningful correlation between the providers and the variance of the outcomes.

question B.3[2] Based on information available in the first three quarters of the initiative, we included Model 2: Orthopedic Surgery Excluding Spine in the study population.

The empirical specification used to estimate the provider-level effects is listed below:

$$Y_{ikt} = \alpha_{kt} + \boldsymbol{X}_{ikt}'\beta + u_{ikt} \qquad\qquad (1)$$

In this example, the dependent variable Y_{ikt} is the dollar value of the payment variable (i.e., total standardized allowed payment for Part A and B services during the anchor hospitalization and the 90-day post-discharge period (PDP) or total standardized allowed payments for Part A institutional services during 90-day PDP, for a given episode i, associated with provider k in time period (i.e., before and after BPCI implementation) t. The vector $\boldsymbol{X}$ includes patient-level characteristics, such as MS-DRG of the qualifying admission, along with the patient's age and HCC indicators, while u_{ikt} represents a random error term. We let the intercept, α_{kt}, vary by provider to allow for provider-level factors to affect the outcome variable, after controlling for patient's characteristics. The provider-level effect, α_{kt}, is a function of market, provider, and program characteristics and period as follows:

$$\alpha_{kt} = \mu_k + \delta_t + \boldsymbol{Z}_{kt}'\theta + v_{kt} \qquad\qquad (2)$$

where μ_k represents a provider fixed effect, δ_t represents the average change in payments from the baseline to the intervention period across all BPCI providers in a given Model and clinical episode group. The vector $\boldsymbol{Z}$ includes a number of indicator variables like the ones in Exhibit 21, that describe the case mix at the provider level, a number of relevant program characteristics (such as gainsharing, SNF waivers, beneficiary incentives), and baseline outcome and market characteristics (SNF beds, market concentration and population in the market above median). Finally, v_{kt} is a provider-level random error.

The advantage of this framework is that it allows us to determine the impact of various provider-level characteristics by risk-adjusting for important patient-level characteristics. The structure of this model is that of a random-coefficient model, allowing us to deal with the potential "within-provider" correlation, as episodes associated with the same provider may potentially be correlated.[33]

---

[32] Descriptive results on risk-adjusted outcomes for clinical groups that did not meet these criteria are in **Appendices R, S, and T.**

[33] To the extent that such correlation does not exist, the multi-level model becomes a simple OLS model with patient-level and provider-level variables. A simple likelihood ratio test helps us determine if OLS or the multi-level model is the optimal approach.

Exhibit 21: Potential Factors Leading to Differences in Pre-post BPCI Changes in Payments across BPCI Providers

Domain	Variables
Patient Characteristics	▪ % of patients between Age 65-79, > Median ▪ % of patients Age 80+, > Median ▪ % of patients Dual Eligible, > Median ▪ % patients Disabled, No ESRD, > Median ▪ % of patients with HCC Case Weight > Median
Program Characteristics	▪ Participating in Gainsharing ▪ Participating in SNF 3-Day Waiver ▪ Participating in Beneficiary Incentives
Provider Characteristics	▪ Facilitator Convener ▪ EI is under AC or DAC ▪ Awardee ▪ FC and Awardee ▪ Prior Bundle ▪ Prior Pay for Performance ▪ Prior Shared Savings ▪ Prior Other ▪ EHR ▪ Health Information Exchange ▪ Beds 0-99 ▪ Beds 100-249 ▪ Beds 250 Plus ▪ Ownership Government ▪ Ownership Nonprofit ▪ Ownership For Profit ▪ % Medicaid Admits >Median ▪ % Medicare Admits >Median ▪ Part A Payment during Baseline >25th%ile ▪ Part A Payment during Baseline >Median ▪ Part A Payment during Baseline >75th%ile ▪ Volume of Episodes during Baseline >25th%ile ▪ Volume of Episodes during Baseline >50th%ile ▪ Volume of Episodes during Baseline >75th%ile ▪ % Medicare Days, >Median ▪ Resident to Bed Ratio, >75th%ile
Market Characteristics	▪ SNF Beds per 10,000 ▪ Herfindahl Index, >Median ▪ Hospital Market Share, >75th %tile ▪ Median Income > Median ▪ Medicare Advantage Penetration >Median ▪ First PAC home >Median ▪ First PAC IRF, > Median ▪ First PAC SNF, > Median ▪ Readmission Rate during Baseline >Median ▪ Disproportionate Share >Median ▪ BPCI Market Penetration-Hospital Level>Median ▪ BPCI Market Penetration-Market Level>Median ▪ Population in Market>Median

Among all potential factors listed in Exhibit 21, we followed three steps to select the variables that were ultimately included in the statistical model. We first included all variables in the model and we calculated the variance inflation factor (VIF) for each variable and for the overall model to determine the degree of collinearity across the variables. The VIF of the initial model, including all variables, was above 45, much higher than a commonly used 'rule-of-thumb' cutoff value of 10. An overall VIF value over 10 indicates that multicollinearity is an issue that leads to the artificial inflation of the standard errors around the regression coefficients (Greene, 2000). Also, the coefficients are unstable (i.e., they change very easily with small changes in the model specification) and are hard to interpret.

We then proceeded to eliminate variables that were highly correlated with each other. We evaluated variables with elevated VIF for exclusion. A correlation matrix including all variables was consulted to confirm which were highly correlated. After successive variable eliminations, we finally achieved a model specification in which the overall VIF was reduced to a value below 11.

Another criterion we used to determine the final model specification was the value of each variable's Pearson coefficient of correlation with the dependent variables. In selecting among highly correlated variables with elevated VIF for inclusion in the model, we gave priority to variables for which Pearson's r was 0.15 or greater, and the p-value (of the Pearson's r statistic) was 0.10 or smaller. **Appendices R, S, and T** present the values of the Pearson's r for each variable considered to be included in the model, by each clinical group and payment outcome considered.

5. Market dynamics analysis

a. Overview

We conducted a preliminary descriptive analysis to determine whether the BPCI program affects non-BPCI hospitals or PAC providers located in BPCI EIs' markets. We hypothesize that Model 2 and 3 EIs would leverage relationships with other types of providers in order to better manage their patients' care across multiple settings, and concentrate their patients' use of these preferred providers to the exclusion of others. For example, Model 2 EIs may seek to discharge patients to certain PAC providers with whom they have gainsharing agreements. Even if there is no gainsharing in place, Model 2 EIs may recommend that their patients use a subset of PAC facilities in their market, potentially those that are more efficient or higher quality. The objective of the analysis is to learn whether the allocation of patients to PAC providers becomes more concentrated in markets with Model 2 BPCI EIs. Likewise, Model 3 EIs may seek to attract patients from hospitals with which they have a good working relationship, or from hospitals that provide better quality of care and better transition planning. This would result in an increased concentration of patients admitted to the PAC EIs from certain hospitals.

We further hypothesize that Model 2, 3, and 4 BPCI EIs would strive to increase their market share of BPCI-eligible episodes, attracting patients that would otherwise go to competitors not participating in BPCI, potentially to increase efficiency, savings, and profits. If so, we would observe an increased volume of BPCI-eligible episodes among EIs relative to other providers of the same types that serve the same kinds of patients in a market.

We created the following measures to examine whether the concentration of patients across PAC providers, or the market share of BPCI EIs, changed over time:

1. "Hospital-to-PACs" concentration indexes for Model 2 EIs

2. "PACs-from-hospital" concentration indexes for Model 3 EIs

3. Share of BPCI-eligible episodes in BPCI EIs ("market share")

We use CBSAs to define BPCI EI's markets. Some CBSAs may be larger than the actual market in which hospitals and PAC providers compete for patients. If the competitive market is not well-defined by CBSA boundaries, we will have limited ability to detect meaningful shifts in market share. For this reason, we examined three very different individual CBSAs that vary by size and density of health care providers, to understand the potential range of effects.

This initial market analysis uses data from the first year of the BPCI program, Q4 2013 through Q3 2014, and two years of pre-BPCI data. Given the short time-frame for observing market-level effects, and the relatively small number of BPCI EIs and patient episodes in the first year, we lack statistical power to detect meaningful changes. The findings should be considered exploratory, and will be extended in future evaluation reports.

The sections below offer more details about our analytic approach.

b. Measure definitions

Hospital-to-PACs concentration index for Model 2 EIs

The Hospital-to-PACs concentration index explores where BPCI patients go after discharge from a Model 2 hospital EI. Separate measures are calculated for patients using SNFs after hospital discharge and those using HHAs. The index is analogous to a Herfindahl Index for PAC care among BPCI patients discharged from Model 2 EIs. Thus, it is defined as the sum of squares of the proportion of beneficiaries that went from a Model 2 hospital EI to each PAC facility. The index can range from zero to one; a near zero index value occurs when small shares of BPCI patients go to a very large number of PAC providers, while the index equals one if all patients discharged from a hospital EI go to a single PAC provider. For example, an increase in the Hospital-to-SNFs concentration index indicates that hospital EI's referrals are going to fewer SNFs than before BPCI. Each patient's first PAC after hospital discharge was included in the calculation. We excluded cases where there was a delay of more than 5 days between hospital discharge and admission to a SNF, or more than 14 days between hospital discharge and admission to an HHA.

We also calculated the highest proportion of the hospital EI's beneficiaries that went to one specific SNF or HHA, as well as the total number of PAC providers who received beneficiaries from the hospital EI. These two measures are complementary, easy to understand ways to identify how relationships strengthen over time between hospital EIs and PAC providers.

PAC-from-Hospitals concentration index for Model 3 EIs

This measure examines the distribution of beneficiaries admitted to a certain PAC, from all the hospitals that treated BPCI-eligible beneficiaries, in a given time period. The measure is calculated separately for Model 3 SNF EIs and for HHA EIs, and is and is defined as the sum of squares of the share of beneficiaries admitted to the PAC EI that were discharged from each

different hospital. An index of zero indicates that small shares of the EI's beneficiaries came from several hospitals, while a value of one indicates that all patients came from a single hospital. An increase in this index indicates that a greater share of a PAC EI's beneficiaries is being referred from fewer hospitals than before. Again, each beneficiary's first PAC after hospital discharge was included in the calculation, and only if the beneficiary was admitted to the PAC (SNF or HHA) within 30 days after hospital discharge.

As with the previous measure, we also calculated the highest proportion of beneficiaries that came from one specific hospital and the total number of hospitals that treated beneficiaries who were subsequently admitted to the PAC EI. These additional measures help to assess the degree to which a Model 3 PAC EI depends on a particular hospital for its BPCI beneficiaries.

Market share of BPCI EIs for BPCI-eligible episodes

An EI's market share is defined as the number of BPCI-eligible episodes admitted to the EI, divided by the total number of the same type of episodes admitted to similar providers (e.g., hospitals, SNFs, HHAs) in the market. We also present the total number of BPCI-eligible episodes for BPCI EIs (i.e., the numerator of the market share measure).

These measures have limitations.   We cannot tell whether the discharges from a hospital to a PAC provider identified in the claims reflect actual referral decision made by the hospital since beneficiaries are free to choose among all PAC providers and we cannot capture specific information about the physician-patient interface using claims. For example, some beneficiaries may arrange for their PAC stay prior to a surgical procedure, based on recommendations from their physician or friends and family.   Some PAC providers use targeted advertising and provide amenities to attract beneficiaries to their facility as a general marketing strategy. The PAC-from-Hospitals concentration index has similar limitations and we cannot determine whether larger shares of beneficiaries come from certain hospitals due to changes in hospital discharge planning, or due to efforts by PAC providers to target and attract certain types of beneficiaries.

c.   *Market definition and selection*

We define the market for a given EI as its CBSA. For Q2 2014, there were a total of 58 CBSAs that contained at least one BPCI EI operating under Model 2 or Model 4. There were 26 CBSAs that contained at least one Model 3 EI (PGPs are excluded from all of the market share analyses).

In addition to reporting the results across all EIs in all markets for each BPCI Model, we also examined a few individual markets for each BPCI Model and EI type. The individual markets reveal a range of results that are influenced by the unique characteristics of the quite different markets we selected. The individual markets selected for case studies are listed in **Exhibit 22**. The markets (CBSAs) were selected based on having a relatively large number of BPCI-eligible episodes and BPCI EIs,[34] and also on having some prior knowledge of the market often based on our site visits.

---

[34] There are two reasons to focus on EIs in markets with a larger BPCI presence. First, we examine trends in the measures across 6 month periods, not only stratifying results by Model and EI type but also stratifying by three

Exhibit 22: CBSAs Selected for Individual Market Analyses

Model	CBSA	High-volume Market	Number of EIs*
2	35644	New York-Wayne-White Plains, NY-NJ	11
2	38060	Phoenix-Mesa-Scottsdale, AZ	7
2	39300	Providence-New Bedford-Fall River, RI-MA	4
3 (SNF only)	16974	Chicago-Naperville-Joliet, IL	11
3 (SNF only)	47644	Warren-Farmington-Hills-Troy, MI	4
3 (HHA only)	16700	Charleston-North Charleston, SC	2
3 (HHA only)	27260	Jacksonville, FL	1
4	36740	Orlando, FL	1
4	28020	Kalamazoo-Portage, MI	1

* This is the number of BPCI EIs participating in at least one of three clinical episode groups: major joint replacement of the lower extremity, congestive heart failure, and sepsis. These episodes represent the largest volume in the first year of the BPCI initiative.

d.   Data

We used 100% of Medicare Part A claims between Q4 2011 and Q3 2014 from providers located in CBSAs where one or more BPCI EIs are located. To construct the concentration indexes we created dyads, or provider pairs, representing any discharging inpatient hospital and the first admitting PAC provider after the beneficiary is discharged. This file consisted of one row per transition from acute to PAC care for an individual patient, starting with the first inpatient claim for a BPCI-eligible episode in Q4 2011 (the start of our pre-BPCI period).

We restricted the data set to BPCI EIs that joined the program during or before Q2 2014 to ensure that each EI had at least one six-month intervention period by Q3 2014 (the last period for which data were available for this analysis). The pre-BPCI period consisted of the 4 six-month periods between Q4 2011 and Q3 2014. The BPCI program allowed rolling entry of EIs over time, therefore a fraction of the BPCI EIs joined in Q4 2013, another larger fraction joined in Q1 2014, and still another fraction joined in Q2 2014 (**Exhibit 23**). This could potentially obfuscate the initial impact of BPCI on the EIs' referral patterns and market shares over the Q4 2013 through Q1 2014 intervention period. We therefore focus on the Q2 2014 through Q3 2014 intervention period – the only six-month intervention period for all BPCI EIs that joined the program during or before Q2 2014 – and demarcate the Q4 2013 through Q1 2014 period to bring attention to the fact that only a fraction of the BPCI EIs included were participating in BPCI during that time. The time period available for analysis was thus relatively brief, and the number of participating EIs was small; both will be expanded in future reports.

---

clinical BPCI episode groups. In markets with a smaller BPCI presence in a given time period, many EIs have few BPCI-related PAC referrals, because most EIs are concentrating on just a few types of BPCI episodes. Insufficient sample sizes tend to generate concentration indexes that vary between the two extreme values, making it difficult to discern any market-specific trends. Second, the behaviors of providers with low volumes of BPCI episodes would likely be less affected by the BPCI program.

Exhibit 23. Time-frame of the Market Dynamics Analysis and Rolling Entry of EIs
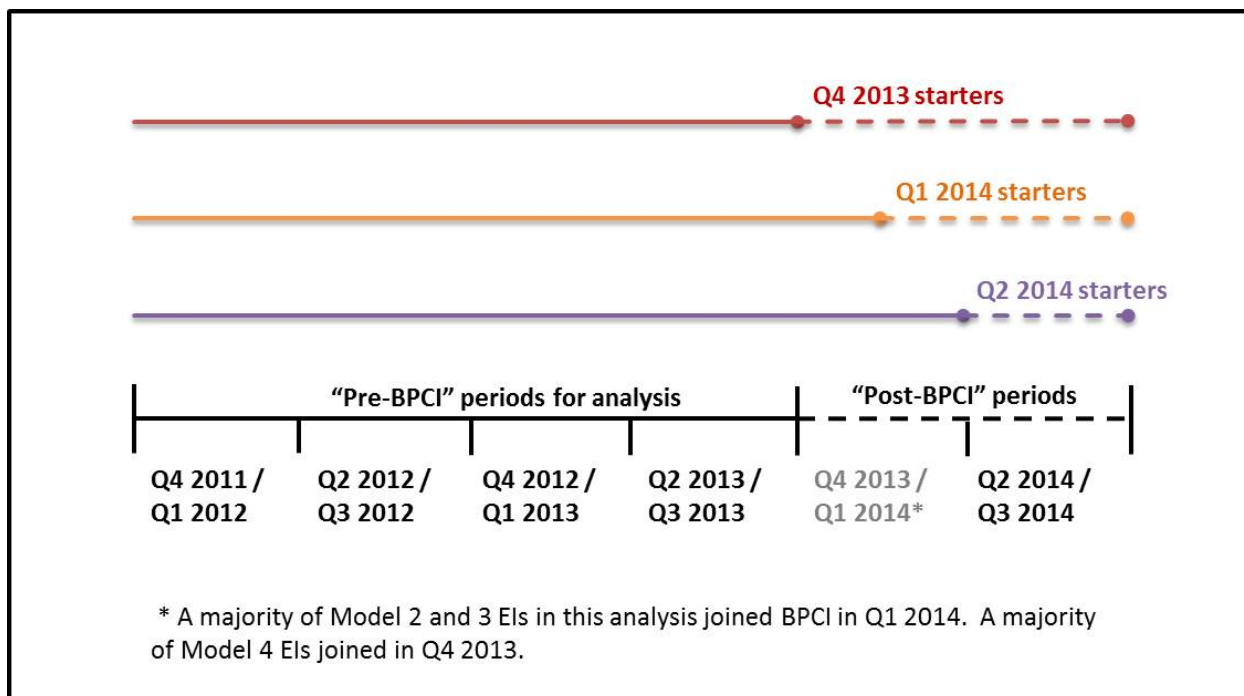
In both overall and market level analyses, we excluded any EIs that did not have at least one admission (for market share) for a hospital-PAC dyad (for concentration indexes) in each of the six time periods, for the clinical episode type being investigated. This step ensures that the entire trend shown pertains to a consistent group of EIs over the full set of baseline and intervention periods. This removes confounding of estimates due to the spurious variation resulting from when providers with relatively low volumes of BPCI-eligible episodes were present in some time periods but not others.

We further restricted the data set to three clinical types of BPCI episodes: major joint replacement of a lower extremity (MJRLE), congestive heart failure (CHF), and sepsis. These are high-volume conditions in which most BPCI EIs were engaged, and include a surgical episode (MJRLE), a chronic medical episode (CHF), and an acute medical episode (sepsis). These three types of clinical conditions represent the largest volume in the first full year of the BPCI initiative.

e. *Analytic approach*

We conducted the market share and referral pattern analysis separately for the three clinical episodes, since market share and referral patterns can be considerably different for different clinical conditions. For example, the preferred PAC for a joint replacement rehabilitation patient may be different than the preferred PAC for a frail CHF patient.

We presented the descriptive trend data for each measure, stratified by type of clinical episode, BPCI Model, and type of EI.[35] Specifically, we calculated the relevant measures for each EI in six-month intervals,[36] from October 2011 to September 2014.

Analyses were conducted at two levels: the overall (nationwide) level and the individual market level. We report the mean and standard deviation of the measures over time periods, using only the sub-sample of EIs located in the market and participating in the relevant Model.

We do not show results pertaining to Physician Group Practice (PGP) EIs because their low BPCI volume in the first year of BPCI would produce spurious variation in the measures and limit the ability to discern reliable trends.

f. Limitations

There are several reasons why this initial, exploratory analysis did not detect many impacts of BPCI on the markets where providers were involved in BPCI. First, it is based on the first year of the initiative, so there was limited time for providers in the market, including BPCI participants, to implement strategic changes in response to BPCI incentives. Second, the rolling entry of EIs could mask changes in EIs' referral patterns and market shares by pooling those varying levels of BPCI experience. Third, the CBSA may not accurately define the local health care markets, particularly for EIs located in large urban areas, such as New York or Chicago. The very small market shares of these EIs suggest that the CBSA is too large to define their local health care market, which makes discerning meaningful changes difficult. Future market analyses will be based on additional time under the initiative, incorporate refined market definitions, and distinguish further among market and provider characteristics.

6. Physician group practice

During the first year of the initiative, there were three Model 2 PGP EIs and one Model 3 PGP EI participating in 90-day orthopedic surgery episodes. Given the small number of PGPs during the first year of the initiative, we did not compare BPCI PGP beneficiaries to a matched comparison group. Instead, we compared PGP risk-adjusted results with all BPCI participants in the same Model and clinical episode group.

Specifically, the risk-adjusted mean outcomes for PGP episodes was compared with the observed mean for all episodes of the same model, clinical episode, and care setting (e.g., initiated in ACH, SNF, or HHA). PGP performance is statistically significantly different than that of all episodes when the observed mean for all episodes lies beyond the upper and lower 95% confidence intervals of the PGP risk-adjusted mean.

---

[35] We did not estimate statistical significance of the difference in measure rates between the baseline and intervention period due to small sample size. Even at the overall nationwide level (all EIs across all markets), the power analysis we conducted suggests the sample size was too small to reliably detect a statistically significant difference if one was present. We will re-evaluate this decision next year.

[36] We used six-month periods instead of quarterly periods to increase the sample size for the measures.

PGP results were risk-adjusted using multivariate regression methods to account for differences in patient mix. The following characteristics were included in the risk-adjustment models:

- MS-DRG corresponding to anchor hospitalization

- Age group and gender

- Medicaid eligibility status

- Medicare eligibility status: disabled (not including ESRD)

- Hierarchical Condition Categories (HCCs) indicators based on diagnostic data from claims during the six months prior to the anchor hospitalization

- Utilization of services in the prior month or prior six months to the anchor hospitalization: acute care hospitalization, SNF, HHA, IRF, LTCH, psychiatric hospital, emergency department

- Census region

We compared the risk-adjusted results for PGPs relative to the observed mean of all episodes for the following samples: Model 2, orthopedic surgery; Model 3, SNF-initiated orthopedic surgery; and Model 3, HHA-initiated orthopedic surgery.

7. *Qualitative analysis*

We conducted an analysis of the qualitative themes from site visits and quarterly interviews to identify themes across case studies, across Awardees, and across markets to understand the range of opinions and experiences related to the BPCI initiative.

For the site visits conducted from October 2014 through September 2015, interviewers developed a single set of notes for the two-day interview session. These notes were then added to the Atlas.ti (version 7.0.91; Scientific Software Development GmbH, Berlin, Germany) qualitative database and coded. The quarterly interviews were recorded and subsequently transcribed. The transcripts for the quarterly interviews were then added to the Atlas.ti qualitative database and coded as well. For both the site visits and quarterly interviews, we conducted two rounds of coding through Atlas.ti. We developed the codes in steps (Crabtree & Miller, 1999), drafting a preliminary code structure after independent review by senior researchers. The first step of coding identified key themes and the second step highlighted recurring and sub-themes. Each document (summary notes for the site visits and transcripts for the quarterly calls) underwent independent review by two researchers with check-in meetings to establish a common understanding and to debrief about what was learned from each site visit and quarterly interview.

We adopted conventional approaches for coding themes, which were based on the questions and characteristics of the BPCI initiative. Further, each person who coded interviews received training in using Atlas.ti and was familiar with the BPCI initiative through program documents, Implementation Protocols, and the evaluation and monitoring plan. Initially, two people coded each set of site visit notes and quarterly interview transcripts to establish a common understanding of how themes would be identified and coded. Themes were developed during the course of coding and recurring themes and sub-themes were coded accordingly.

Analysis of the qualitative themes from case studies and quarterly interviews was guided by Research Questions A and C and the constant comparative method, a systematic data coding and analysis process (Glaser & Strauss, 1967) during which specific quotes were categorized into themes with codes developed iteratively to reflect the data. After coding transcripts, we reviewed the themes relevant to the specified research questions, cataloguing the themes by specific topics (e.g., organizational structure, waiver use, PAC utilization). Coding results were compared to identify concordant themes. Discrepancies were discussed until final consensus was reached. This process occurred until saturation in the coding was achieved. As needed, we added new codes to capture new concepts. After finalizing the code list, a team of coders worked through all case study summary notes and quarterly interview and focus group transcripts.

III. Model 2 Results

This section presents information about the experience of Model 2 BPCI participants and their episodes of care, organized by research question. Quantitative analyses of Medicare claims data are based on episodes initiated during the first full year of BPCI (Q4 2013 to Q3 2014); patient assessment data are based on episodes initiated during the first three quarters of BPCI (Q4 2013 to Q2 2014); Awardee-submitted data are based on calendar year 2014 or Q4 2014 depending on the measure; and survey data are based on episodes initiated during June, October, and November 2014. Qualitative data from interviews reflect the first six quarters of the BPCI initiative (Q4 2013 to Q12015) while site visits results reflect the first seven quarters of the BPCI initiative (Q4 2013 to Q2 2015). The quantitative outcomes are risk-adjusted as described in Section II.D.2 above.

A. Characteristics of the Initiative and Participants

1. Participants

This section describes BPCI Model 2 participants that joined during the first year of the initiative (Q4 2013 through Q3 2014). By the third quarter of 2014, 61 participants with 110 hospital Episode Initiators (EIs) and 3 physician group practices (PGPs) were active in Model 2 of BPCI. See **Appendix Q** for a detailed description of the growth of the initiative through July 2015.

Exhibit 24 compares first year BPCI-participating hospital EIs to non-participating hospitals. Medicare inpatient days represented a smaller proportion of total inpatient days among the participants (37%) than non-participants (41%). Disproportionate share percent was similar between BPCI participants (30%) and non-participants (28%), which indicates that both groups reach about the same proportion of Medicare Supplemental Security Income, Medicaid, or other low-income populations. Additionally, the proportion of BPCI-participating EIs that were part of a chain was comparable to that of non-participating hospitals (49% vs. 52%).

BPCI-participating hospital EIs more often were non-profit, set in urban locations, and had a higher bed count than hospitals that did not participate in BPCI during the first year of the initiative. A larger proportion (85%) of participating hospitals was non-profit, compared with 59% of non-participating hospitals; participating EIs were less often government facilities (4% vs. 18%) or for-profit hospitals (12% vs. 22%). Geographically, nearly all of BPCI-participating hospitals were in urban locations (95%), compared with 71% of non-participating hospitals. BPCI-participating hospitals had about twice the average bed count (359 vs. 188). They had a higher resident-to-bed ratio than non-participating hospitals (0.18 vs. 0.06) and had nearly twice as many admissions for BPCI episode MS-DRGs during 2011 (4,060 vs. 2,140).

BPCI-participating hospital EI episodes had higher standardized payments for the inpatient stay plus the 90 day PDP in 2011 relative to episodes from non-participating hospitals. The difference in standardized payments varied by clinical episode group; standardized payments were 0.5% higher for cardiovascular surgery episodes initiated in BPCI participating hospitals, for example, and 7% higher for non-surgical other medical episodes.

Exhibit 24: Characteristics of BPCI-participating Hospital EIs and Non-participating Hospitals, Model 2, Q4 2013 - Q3 2014

	BPCI Hospital EIs (N=110)		Non-participating Hospitals (N=3,056)	
	N	**%**	**N**	**%**
Ownership				
For Profit	13	12%	672	22%
Government	4	4%	550	18%
Non-Profit	93	85%	1,803	59%
Urban/Rural				
Rural	6	5%	886	29%
Urban	104	95%	2,170	71%
Part of Chain				
Yes	54	49%	1,528	52%
	Mean		**Mean**	
Bed Count	359		188	
Number of Admissions for BPCI Episode MS-DRGs, 2011	4,060		2,140	
Medicare Days Percent	37%		41%	
Resident-to-bed ratio	0.18		0.06	
Disproportionate Share Percent	30%		28%	
Standardized Part A Allowed Payment inpatient stay plus 90 day PDP, 2011				
Clinical Episode Group (N, BPCI discharges)	**Mean**		**Mean**	
Orthopedic surgery (38,718)	$29,439		$28,882	
Non-surgical other medical (14,577)	$26,595		$24,897	
Non-surgical neuro (2,690)	$25,235		$24,040	
Non-surgical respiratory (19,818)	$24,310		$23,183	
Non-surgical cardiac (30,290)	$22,191		$21,677	
Surgical and non-surgical GI (5,854)	$21,935		$20,835	
Cardiovascular surgery (18,452)	$33,664		$33,513	
Spinal surgery (2,504)	$29,781		$28,929	

Source: Lewin analysis of 2013 POS and 2011 Medicare claims. BPCI participating hospitals are defined as Model 2 EIs, Q4 2013 – Q3 2014. Non-participating hospitals are all other hospitals and exclude Model 4 hospitals participating in BPCI during the first year.

a. *Awardee submitted baseline characteristics*

Beginning in Q1 2015, Awardees and EIs submitted information about their experiences that may have contributed to their implementation of BPCI through an on-line data reporting tool. This included patient payer mix, prior experience with care redesign and payment initiatives, and quality of care indicators. We collected data representing the Awardee and EI experience during various time periods, including prior to joining BPCI and during 2014.

Patient payer mix

Payer mix was reported for 99 Model 2 EIs (96 hospitals and 3 PGPs), based on the number of patients in 2014 with a primary payer of Medicaid, Medicare, private, or other. As shown in Exhibit 25, Medicare patients and private payer patients comprised the highest portion of total patients for Model 2 EIs, with approximately equal average shares, although the distribution of Medicare and private patients varied widely across hospitals.

Exhibit 25: Patient Payer Mix, Model 2, 2014

	BPCI Model 2 EIs N=99			
	Mean	**Min**	**Median**	**Max**
Percentage of Medicaid Patients	18.2%	0.0%	16.6%	78.4%
Percentage of Medicare Patients	34.9%	2.7%	34.4%	95.2%
Percentage of Private Payer Patients	34.5%	0.0%	32.3%	87.4%
Percentage of Other Payer Patients	12.4%	0.0%	7.2%	75.0%

Source: Lewin analysis of Awardee-submitted data collected February 2015 for Model 2 EIs participating in BPCI between Q4 2013 – Q4 2014.

Prior experience

As summarized in Exhibit 26, Awardees reported whether their Model 2 EIs participated in care redesign or payment incentive initiatives prior to their participation in BPCI. Care coordination was the most commonly reported among Model 2 EIs, with 71.3% reporting prior participation. Experience in redesign of care pathways, enhancement in care delivery, as well as patient activation, engagement and risk management were also commonly reported among Model 2 EIs.

The most common type of payment incentives experience among Model 2 participants was pay for performance, with 67.3% of EIs reporting prior participation. Close to half of the participants also reported prior experience in shared savings (48.5%). Only 16.8% of EIs reported prior experience with bundled payments.

Exhibit 26: Care Redesign and Payment Incentives Experience, Model 2, Prior to BPCI Participation

		Model 2 EIs N=101	
		N	%
Prior experience in care redesign initiatives:	Redesign of Care Pathways	63	62.4%
	Enhancements in Care Delivery	65	64.4%
	Patient Activation, Engagement & Risk Management	63	62.4%
	Care Coordination	72	71.3%
	System Changes to Support Care	51	50.5%
	Other Redesign Activities	10	9.9%
Prior experience in payment incentives:	Bundled Payments	17	16.8%
	Pay for Performance	68	67.3%
	Shared Savings	49	48.5%
	Other Payment Incentives	18	17.8%

Source: Lewin analysis of Awardee-submitted data collected February 2015 for Model 2 EIs participating in BPCI between Q4 2013 – Q4 2014.

Electronic health record (EHR) use

Model 2 EIs have high rates of EHR use, as shown in Exhibit 27. Of the 101 EIs in Model 2, 100 reported using an EHR system. The EIs that use an EHR system were asked to report the meaningful-use functionalities of their systems. Computerized physician order entry was the most common functionality reported by Model 2 EIs (99%), followed by discharge instructions and care summary documents (97%). Medication management and clinical decision support were also reported by 95% and 90% of Model 2 EIs, respectively.

Exhibit 27: Electronic Health Record (EHR) Use, Model 2, 2014

		Model 2 EIs N=101	
		N	%
	EIs with an EHR	100	99.0%
Meaningful-use functionalities (among EIs with EHRs):	**Automated Quality Reporting**	78	78.0%
	Discharge Instructions and Care Summary Documents	97	97.0%
	Medication Management	95	95.0%
	e-Prescribing	72	72.0%
	Computerized Physician Order Entry	99	99.0%
	Clinical Decision Support	90	90.0%

Source: Lewin analysis of Awardee-submitted data collected February 2015 for Model 2 EIs participating in BPCI between Q4 2013 – Q4 2014.

EIs that use EHRs were also asked to report the health information exchange (HIE) capabilities of their EHR systems, as shown in Exhibit 28. Of the 100 Model 2 EIs with EHR systems, 93% reported having HIE capabilities. The majority of these EIs report using these capabilities to

exchange information with other health care providers (87.1%). Use of HIE capabilities to provide information to patients was less common among EIs (66.7%).

Exhibit 28: Health Information Exchange (HIE) Capabilities among Episode Initiators with EHRs, Model 2, 2014

		Model 2 EIs N=100	
		N	%
	EIs using EHRs that have Health Information Exchange Capabilities	93	93.0%
Use of HIE capabilities (among those with such capabilities) to exchange data with:	**Providers**	81	87.1%
	Patients	62	66.7%
	None of the Above	8	8.6%

Source: Lewin analysis of Awardee-submitted data collected February 2015 for Model 2 EIs participating in BPCI between Q4 2013 – Q4 2014.

The eight Model 2 EIs that reported that they did not have HIE capabilities (n=7) or an EHR system (n=1), were asked to report the alternative methods they use to exchange information with providers and patients. As illustrated in Exhibit 29, 100% of these EIs reported that they exchange information with providers via fax and 75% reported using the telephone. These methods were less commonly used to exchange information with patients, with 12.5% using fax and 37.5% using the telephone to communicate information with patients. Three quarters of these EIs reported using "other" methods for exchanging information with patients. Such methods include mail and patient portals.

Exhibit 29: Methods of Exchanging Information with Providers and Patients among Episode Initiators without EHRs or HIE Capabilities, Model 2, 2014

		Model 2 EIs N=8	
		N	%
Methods of exchanging information with providers:	**Telephone**	6	75.0%
	Email	5	62.5%
	Fax	8	100.0%
	other methods	5	62.5%
	Do not exchange information	0	0.0%
Methods of exchanging information with patients:	**Telephone**	3	37.5%
	Email	0	0.0%
	Fax	1	12.5%
	other methods	6	75.0%
	Do not exchange information	1	12.5%

Source: Lewin analysis of Awardee-submitted data collected February 2015 for Model 2 EIs participating in BPCI between Q4 2013 – Q4 2014.

b. Awardee submitted quality measures

Awardees and EIs also submitted information about certain quality measures through the on-line data reporting tool beginning in Q1 2015.

Medication reconciliation at discharge

Model 2 EIs reported the number of BPCI patient discharges among patients 65 and older for whom medications were reconciled at discharge and the total number of discharges among this population during Q4 2014. As displayed in Exhibit 30, the mean proportion of BPCI beneficiaries who received medication reconciliation at discharge was 92% across all Model 2 EIs, and the majority reported that every BPCI beneficiary who was discharged received medication reconciliation.

Exhibit 30: Proportion of BPCI Beneficiaries Age 65+ that received Medication Reconciliation at Discharge, Model 2, Q4 2014

	Model 2 EIs (N=99)			
	Mean	**Median**	**Minimum**	**Maximum**
Proportion of BPCI beneficiaries age 65+ that received medication reconciliation at discharge	92%	100%	0%	100%

Source: Lewin analysis of Awardee-submitted data collected February 2015 for Model 2 EIs participating in BPCI during Q4 2014.

FDA reportable events

Model 2 EIs report any BPCI beneficiary deaths or serious injuries that are reportable to the FDA that occurred during Q4 2014. As displayed in Exhibit 31, no EIs reported any adverse events associated with the use of contaminated drugs, devices, or biologics provided by the health care setting. Only one EI reported adverse events associated with the use or function of a device in patient care, in which the device was used or functioned other than as intended.

Exhibit 31: Proportion of BPCI Beneficiaries that Experienced a Patient Death or Serious Injury Reportable to the FDA, Model 2, Q4 2014

	Patient Death or Serious Injuries Reportable to the FDA by Model 2 EIs (N=99)			
	Mean	**Median**	**Minimum**	**Maximum**
Proportion of BPCI beneficiaries that experienced a patient death/injury associated with the use of the contaminated drugs, devices, or biologics provided by the health care setting	0.000%	0.000%	0.000%	0.000%
Proportion of BPCI beneficiaries with an injury due to the use/function of a device in patient care, in which the device is used or functions other than as intended	0.002%	0.000%	0.000%	0.213%

Source: Lewin analysis of Awardee-submitted data collected February 2015 for Model 2 EIs participating in BPCI during Q4 2014.

2. Market characteristics

The 110 Model 2 BPCI-participating hospitals and 3 PGPs and 20 Model 4 BPCI-participating hospitals were located throughout the country, as pictured in Exhibits 32 and 33. In this section, we present the market characteristics[37] of the hospitals and PGPs that participate in Models 2 and 4 (BPCI markets) and the characteristics of markets with no BPCI-participating hospitals or PGPs (non-BPCI markets) through Q3 2014.[38] It should be noted that the non-BPCI markets include those markets with EIs in Phase 1 of the program. Thus, some non-BPCI markets will become BPCI markets with the large influx of providers that transition to Phase 2 in Q2 and Q3 of 2015.

Exhibit 32: Number of BPCI-Participating Hospitals by CBSA, Model 2 and Model 4, Q4 2013 – Q3 2014
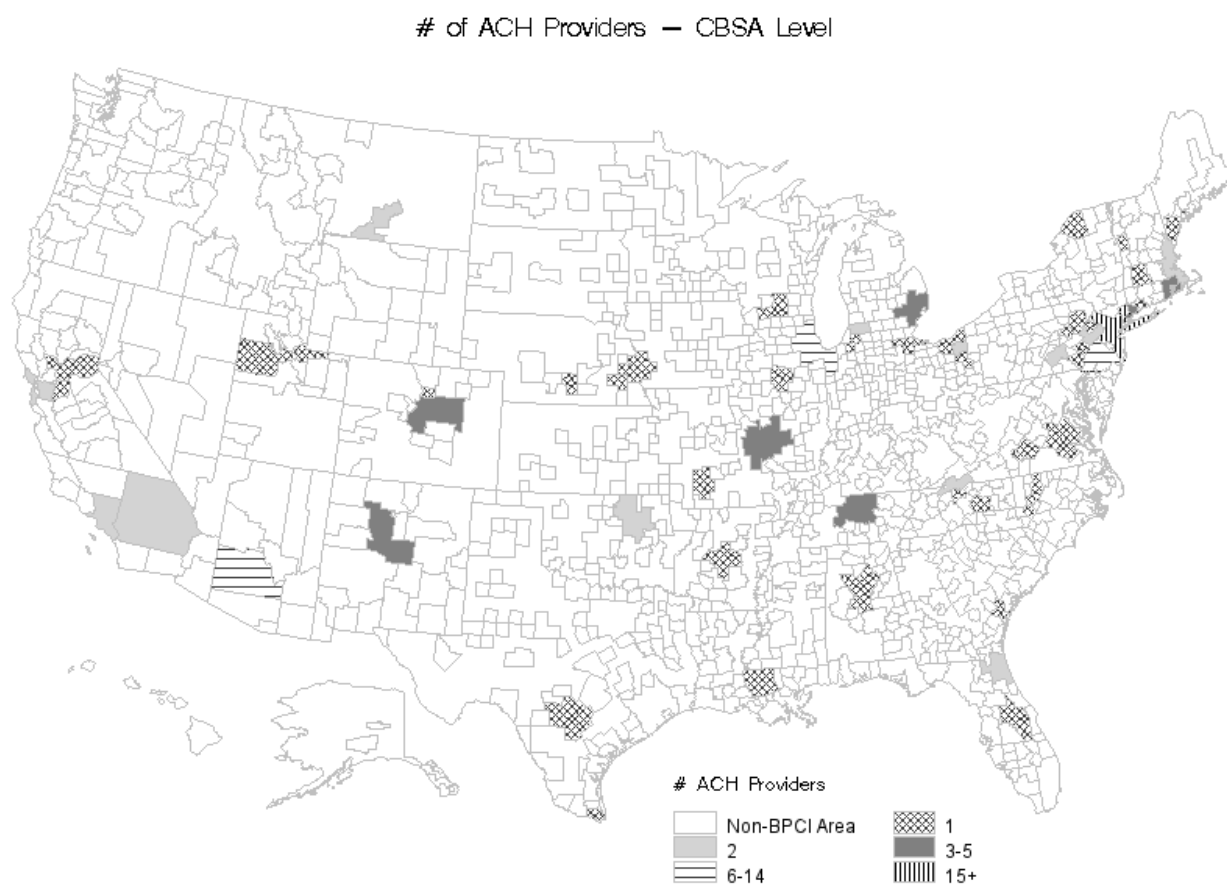
Source: Lewin analysis of CMS' BPCI database for all Q4 2013 – Q3 2014 BPCI participating hospital EIs.

---

[37] The market is defined as the Core Based Statistical Area (CBSA). Providers not located in a CBSA were assigned to the largest CBSA within their Hospital Referral Region (HRR).

[38] Non-BPCI markets are CBSAs that do not have a Model 2 or 4 BPCI participant. Areas of the country that are not in a CBSA are therefore not included in these non-BPCI markets.

Exhibit 33: Number of BPCI-Participating PGPs by CBSA, Model 2, Q4 2013 – Q3 2014
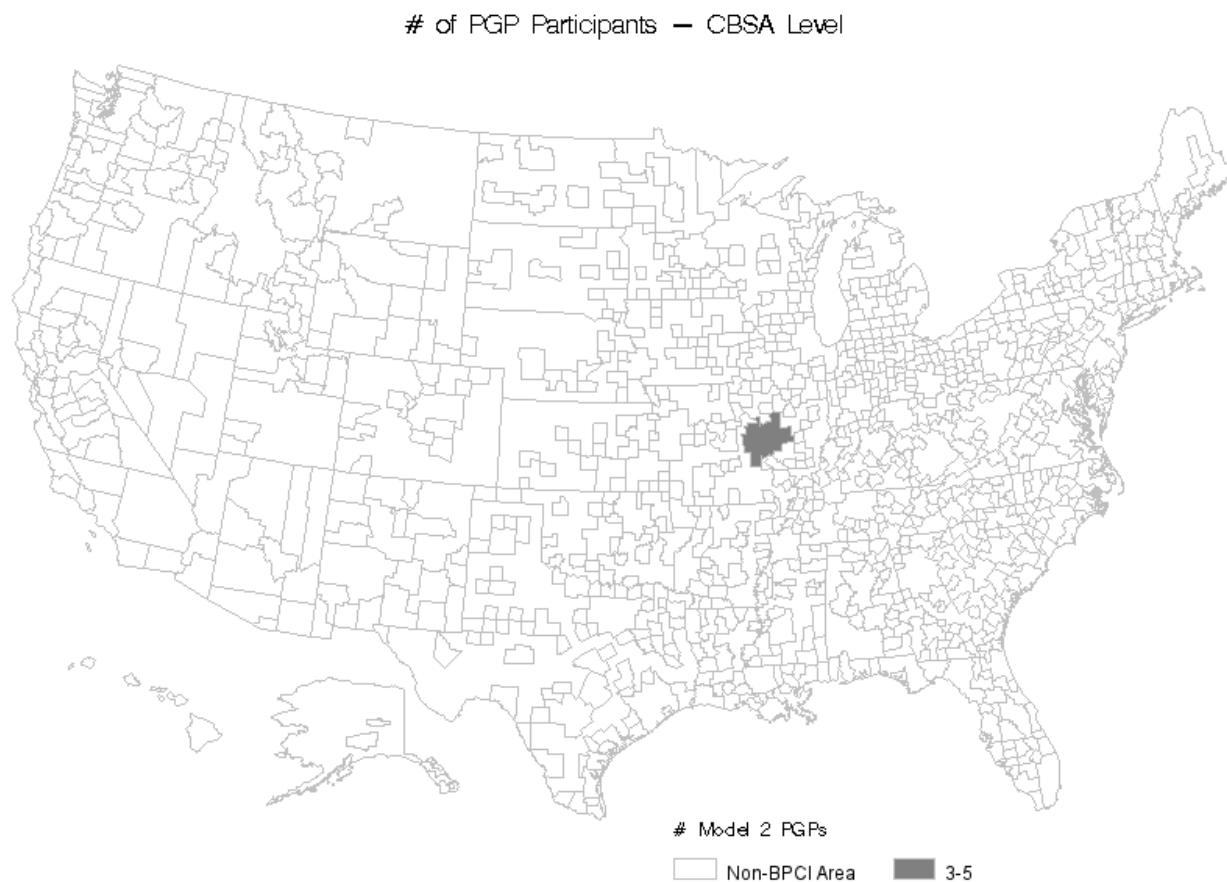
# of PGP Participants — CBSA Level

Model 2 PGPs

☐ Non-BPCI Area   ■ 3-5

Source: Lewin analysis of CMS' BPCI database for all Q4 2013 – Q3 2014 BPCI participating PGP EIs.

Markets with BPCI-participating hospitals differed from markets without BPCI hospitals. As shown in Exhibit 34, BPCI markets tended to have larger populations (average 1.8 million residents), whereas non-BPCI markets were smaller (average 206,000 residents). On average, median household income was higher in BPCI markets ($51,100) than in non-BPCI markets ($43,700). BPCI markets had, on average, higher Medicare Advantage (MA) penetration than non-BPCI markets (26.9% vs. 17.6%).

Consistent with location in larger, more urban areas, there was a higher concentration of primary care physicians (PCPs) per 10,000 residents in the BPCI markets (8.2 vs. 6.3), as well as a higher concentration of physician assistants/nurse practitioners (PA/NPs) per 10,000 residents (7.8 vs. 6). The difference between BPCI and non-BPCI markets was more pronounced with respect to specialists per 10,000 residents (11.2 vs. 5.1). BPCI markets had fewer SNF beds (58.2 vs. 71.4) per 10,000 residents than non-BPCI markets. The proportion of Medicare-aged residents was similar for BPCI and non-BPCI markets. On average, 14% of residents in BPCI markets were 65 or older, relative to 15% in non-BPCI markets.

First year Model 2 and Model 4 BPCI-participating hospitals tended to be in CBSAs that were more competitive than non-BPCI markets. The Herfindahl index, defined as the sum of the

squares of the market shares within a provider type, was used in assessing market competition. Higher Herfindahl index values generally indicate lower competition and greater market power for local providers. The relatively low mean Herfindahl index for hospitals in BPCI markets (0.29) suggests a high degree of competition among hospitals, with none of them dominating the market. In contrast, the mean Herfindahl index in non-BPCI markets was higher (0.69), suggesting that these markets were relatively less competitive and probably dominated by fewer hospitals with more market share.

Exhibit 34: Characteristics of BPCI Markets and Non-BPCI Markets, Model 2 and Model 4, Q4 2013 – Q3 2014

Market Characteristics Models 2 & 4	BPCI Markets N=60; 6.4% of Markets				Non-BPCI Markets N=882; 93.6% of Markets			
	Mean	Median	25th	75th	Mean	Median	25th	75th
Hospital Penetration	38.5%	31.5%	16.0%	46.5%	0.0%	0.0%	0.0%	0.0%
Herfindahl Index – hospital	0.29	0.25	0.11	0.34	0.69	0.95	0.43	1.00
Herfindahl Index – SNF	0.07	0.04	0.02	0.07	0.33	0.28	0.15	0.41
Herfindahl Index – HHA	0.22	0.14	0.07	0.27	0.54	0.52	0.23	1.00
Herfindahl Index – IRF	0.38	0.26	0.00	1.00	0.11	0.00	0.00	0.00
Medicare Advantage Penetration	26.9%	25.0%	17.9%	36.9%	17.6%	14.7%	8.3%	23.7%
Population	1,830,486	698,835	360,072	1,842,713	206,728	67,698	38,885	150,811
Median Household Income	$51,069	$50,101	$46,222	$55,076	$43,741	$42,570	$38,276	$48,029
% Age 65+	14%	13%	12%	15%	15%	15%	13%	17%
PCPs Per 10,000	8.2	8.0	7.3	8.8	6.3	6.1	4.7	7.5
Specialists Per 10,000	11.2	10.2	7.5	13.1	5.1	4.3	2.5	6.5
PA/NPs Per 10,000	7.8	7.2	5.5	9.1	6.0	5.5	3.8	7.5
SNF Beds Per 10,000	58.2	57.8	40.8	75.0	71.4	65.1	43.8	91.5
LTCH Beds Per 10,000	1.1	0.8	0.3	1.6	0.5	0.0	0.0	0.0
IRF Beds Per 10,000	0.7	0.2	0.0	1.1	0.3	0.0	0.0	0.0
CAH Beds Per 10,000	0.4	0.0	0.0	0.6	1.8	0.0	0.0	1.5

Source: Lewin analysis of 2011 Medicare claims and 2011 AHRF.

Notes: "Hospital Penetration is the percentage of Medicare admissions in the 48 clinical episodes in the market attributed to the BPCI-participating hospitals in the market. Variable definitions are in **Appendix N**

3. Model incentive structure characteristics

a. Entry decisions

During the first four quarters of the initiative, 61 of the 94 Awardees participated in Model 2, accounting for 113 of the 227 EIs (110 hospitals and 3 PGPs). Model 2 also accounted for the largest share of episodes initiated during Q4 2013 to Q3 2014 at 42,572, or 73% of all episodes observed across the three models.

The Awardee interviews and EI case studies provided insights into the reasons that organizations chose to participate in BPCI. We conducted 11 case studies and 57 Awardee interviews with Model 2 participants. During the first year of the initiative, approximately half of the respondents said that the BPCI initiative provided an opportunity to learn about bundled payments and to experiment with new payment models. This topic also arose during several of the Model 2 case studies conducted during this year. More than half of the Model 2 respondents we interviewed claimed that the learning opportunities were the main reason for their interest in BPCI, making this the most common reasoning heard by the research team.

> *"Bundled payments are the way of the future, so we wanted to figure it out now."*

In addition, a smaller number of EIs indicated that they joined BPCI after being approached by either an Awardee Convener (AC) or Facilitator Convener (FC). One hospital reported that they were contacted by three different conveners but otherwise did not know about the initiative.

Model 2 Awardees also considered risk and opportunity when making choices about the structure of BPCI design. Interviewees frequently cited these factors in their decisions regarding model, episode, and episode length.

According to the Awardees we interviewed, Model 2 was selected because the bundles include the hospital stay and post-acute care (PAC). Many Awardees indicated that managing PAC offers the greatest opportunity for achieving savings. This rationale was also cited during case studies conducted this year; participants reported selecting Model 2 because of the perceived financial opportunity associated with PAC. One participant noted, however, that they were considering switching to Model 4 so that they would not have to worry about how they would influence and reduce PAC utilization. Participants that chose Model 2 also indicated that they wanted to continue receiving retrospective, fee-for-service payments, citing the difficulty of adapting to prospective payments because of the increased risk and need for infrastructure changes (e.g., revamped internal accounting systems). During a case study held in 2015, one EI noted that Model 2 was selected specifically to minimize financial risk, which they felt would have been greater under Model 4. Finally, one EI highlighted CMS' influence in their decision; when they were applying, they were not sure which model to select, and CMS encouraged them to choose Model 2.

b. Episode and length selection

The count of EIs participating in each of the 48 clinical episodes during the first four quarters of the initiative is shown in Exhibit 35.   Model 2 EIs participated in all 48 clinical episodes during Q4 2013 to Q3 2014. The average Model 2 EI participated in 5 clinical episodes. Orthopedic surgery was the most popular clinical group, with 75% of EIs selecting at least one clinical episode within this group. Nearly three-quarters of EIs participated in major joint replacement of the lower extremity, making it the most popular clinical episode among Model 2 participants. Congestive

heart failure was the next most common clinical episode, chosen by 35% of EIs. During the first four quarters of the initiative, EI participation was less than 10% in 28 clinical episodes. None of the EIs participated in all 48 clinical episodes.

Exhibit 35: Participation of Episode Initiators by Clinical Episode, Model 2, Q4 2013 - Q3 2014

Clinical Episode	Episode Initiators by Participant Type (N=113)		
	ACH (N=110)	PGP (N=3)	%
Non-surgical and surgical: Gastrointestinal (GI)			
Esophagitis, gastroenteritis and other digestive disorders	5	0	4%
Gastrointestinal hemorrhage	10	0	9%
Gastrointestinal obstruction	7	0	6%
Major bowel procedure	10	0	9%
Total	**17**	**0**	**15%**
Non-surgical: Cardiovascular			
Acute myocardial infarction	17	0	15%
Atherosclerosis	12	0	11%
Cardiac arrhythmia	9	0	8%
Chest pain	8	0	7%
Congestive heart failure	39	0	35%
Medical peripheral vascular disorders	10	0	9%
Syncope & collapse	5	0	4%
Total	**41**	**0**	**36%**
Non-surgical Neurovascular			
Stroke	13	0	12%
Transient ischemia	5	0	4%
Total	**15**	**0**	**13%**
Non-surgical Orthopedic			
Fractures of the femur and hip or pelvis	12	0	11%
Medical non-infectious orthopedic	12	0	11%
Total	**15**	**0**	**13%**
Non-surgical: Other Medical			
Cellulitis	10	0	9%
Diabetes	7	0	6%
Nutritional and metabolic disorders	5	0	4%
Red blood cell disorders	5	0	4%
Renal failure	7	0	6%
Sepsis	13	0	12%
Urinary tract infection	5	0	4%
Total	**16**	**0**	**14%**

Clinical Episode	Episode Initiators by Participant Type (N=113)		
	ACH (N=110)	PGP (N=3)	%
Non-surgical: Respiratory			
Chronic obstructive pulmonary disease, bronchitis, asthma	29	0	26%
Other respiratory	5	0	4%
Simple pneumonia and respiratory infections	23	0	20%
Total	**33**	**0**	**29%**
Cardiovascular surgery			
AICD generator or lead	3	0	3%
Cardiac defibrillator	6	0	5%
Cardiac valve	12	0	11%
Coronary artery bypass graft	17	0	15%
Major cardiovascular procedure	9	0	8%
Other vascular surgery	10	0	9%
Pacemaker	9	0	8%
Pacemaker device replacement or revision	4	0	4%
Percutaneous coronary intervention	8	0	7%
Total	**30**	**0**	**27%**
Orthopedic surgery			
Amputation	5	0	4%
Double joint replacement of the lower extremity	13	0	12%
Hip & femur procedures except major joint	20	0	18%
Lower extremity and humerus procedure except hip, foot, femur	14	0	12%
Major joint replacement of the lower extremity	81	3	74%
Major joint replacement of the upper extremity	12	2	12%
Other knee procedures	1	0	1%
Removal of orthopedic devices	13	0	12%
Revision of the hip or knee	18	0	16%
Total	**82**	**3**	**75%**
Spinal surgery			
Back & neck except spinal fusion	5	0	4%
Cervical spinal fusion	11	0	10%
Combined anterior posterior spinal fusion	9	0	8%
Complex non-cervical spinal fusion	9	0	8%
Spinal fusion (non-cervical)	12	0	11%
Total	**20**	**0**	**18%**

Source: Lewin Analysis of CMS' BPCI database, June 2015.

Note: The total number of EIs for a given clinical group will not add up to the total EIs participating in each of the clinical episodes within that group because EIs can participate in more than one clinical episode within the group.

Awardees interviewed in both the first and second years of the initiative indicated that the decisions about BPCI episode selection were typically made jointly by the hospitals' administration (e.g., CEO, CFO) and clinical leadership (e.g., chief medical officer, physician champions).

Awardees' reasons for selecting particular episodes included a number of business considerations. Major joint replacement of the lower extremity continued to be the most commonly selected episode because it is typically an elective and relatively predictable procedure with less variable outcomes. Participants reported they could usually prepare patients before these surgeries, which would not be possible for unplanned episodes. As interviews progressed into the second year, a common concern among Awardees was that the joint replacement episodes included joint replacements following fractures, which are very different from elective joint replacements. For example, patients who received joint replacements following fractures were likelier to require post-acute care and typically had longer lengths of stay. Furthermore, though care redesign often emphasized pre-operative planning and patient education before joint replacement surgery, these interventions could not be applied to fracture patients.

Participants indicated that their existing expertise, physician engagement, learning opportunities, savings opportunities, and patient volume were all important considerations in episode selection. Several participants indicated that they selected episodes for which they had strong relationships with physicians. The opportunity for

> *"One of the reasons why we selected CHF as our program for the bundle was that, first, we knew we had extremely high quality physicians that we'd be working with and secondly, we knew that there would be a good partnership relationship with these docs."*

learning and education was frequently cited as an important factor in episode selection. While most participants said they selected joint replacement episodes to learn about BPCI while not taking on too much financial risk, two participants specifically selected the pneumonia episode because it provided an opportunity to learn about managing more medically complicated patient populations. Awardees noted in the interviews that they selected episodes based on opportunities for savings or for quality improvement. Finally, several participants selected episodes partly based on whether they had sufficient patient volume.

Through Q3 2014, Model 2 participants participated in 30-, 60-, and 90-day episodes. More than three quarters of the Model 2 participants selected 90-day episodes (82%); fewer participants chose 30-day episodes (15%) and 60-day episodes (2.7%). According to the Awardee quarterly interviews, choice of episode length was primarily based on the ability to enhance opportunities or mitigate risk.

The Awardees that chose the 30-day episodes indicated they selected the shorter window because it allowed for greater risk control. These Awardees believed that a longer period was riskier because they felt they would have to assume financial risk for too many unrelated complications that could lead to hospital readmissions.

Participants that chose the 90-day episode length specified three major reasons for selecting this duration: the desire to control a longer continuum of care, the cost reduction opportunities in PAC, and the lower discount rate (applied to the target amount) for the 90-day period. One participant indicated that they were considering a switch from 90-day to 30-day episodes to reduce their risk, as managing the 90-day period has proved challenging.

Participants also indicated that they selected episode length based on what was clinically appropriate for the episode. For example, respondents noted that 90 days was more appropriate for a chronic condition such as chronic obstructive pulmonary disease (COPD). Other respondents indicated that 30 days was most appropriate for joint replacements, although 90 days was by far the most common episode length among Model 2 participants that selected this episode.

According to quarterly calls and case studies, several participants encountered challenges with respect to the way BPCI episodes are defined. Awardees and EIs indicated that episodes may include services that the hospital has no control over, such as mental health or substance abuse services, or that pre-planned services which the hospital feels are clinically appropriate are nonetheless designated as inappropriate readmissions under BPCI. Another challenge identified by participants was the transfer of a beneficiary to Medicare Advantage; in these instances, the beneficiary was no longer considered a BPCI patient, which impacted reimbursement.

c. *Conveners in BPCI*

When entering the BPCI initiative, participants select one of several roles. The participant's role indicates whether it bears risk with CMS, initiates episodes, or serves solely an administrative function (e.g., as a non-episode initiating convener). There are three different types of convener roles in BPCI: FC, DAC, and AC. FCs are non-episode initiating and non-risk bearing participants. These participants typically serve administrative and technical assistance functions on behalf of DACs and DAs. ACs and DACs function similarly, with the exception that DACs participate in BPCI under an FC. ACs and DACs may initiate episodes, but more notably, they assume financial risk on behalf of their EIs.

From Q4 2013 through Q3 2014, there were five FCs and 20 ACs participating in Model 2. Almost one-third of Model 2 EIs were participating under two large ACs, accounting for over 50% of episodes initiated during the first year of the initiative.

To better understand the role of these conveners in BPCI, one set of quarterly interviews conducted this past year focused on FCs and their reasons for joining the BPCI initiative in that role. Representatives from four different Model 2 FCs participated in these interviews.

According to those interviewed, FCs were largely motivated by a desire to increase their expertise with alternative payment models. When discussing their decision to participate as an FC, two interviewees cited an aversion to financial risk. The FCs interviewed believed that their primary role in BPCI was to educate their participating EIs. These participants expressed a desire to demonstrate how care processes could be redesigned in a way that both improved quality and reduced costs.

> *"We're a not-for-profit organization. We would never think about taking risk or accepting savings or sharing savings with folks. So there was never any question about us being a Facilitator Convener."*

As discussed in the interviews, FCs typically provided contractor-related support with regards to BPCI. These services included general education, program administration, and data analytics. Although some FCs indicated that they made recommendations about episode and waiver selection based on the data analysis, other FCs did not attempt to influence their participants' decisions. Some FCs also provided advocacy support and acted as the intermediary between participants and CMS.

> *"I would say that probably one of the smartest things that CMS did was permit the facilitators to be part of this program because, at least my observation, it is just too heavy of a lift for individual hospitals to both undertake the care redesign that's necessary over the long run, as well as understand and interpret all of the data and the policies of the program."*

The role of an FC differs from that of an AC or DAC. During site visits, Model 2 EIs described the varying roles and responsibilities of their ACs and DACs under the BPCI initiative. ACs and DACs also explained the role they play on behalf of their EIs during Awardee interviews. These discussions proved that no one convener structure is alike; while some ACs and DACs play active roles in the administration and implementation of BPCI, others adopt a more passive approach, providing resources and guidance.

Two key areas of convener involvement that participants described include their role in episode selection and in implementing care redesign. When speaking to their AC's involvement in episode selection, one EI explained that its AC encouraged them to select a large number of episodes in order to increase the potential for greater savings under BPCI. Ultimately the EI disagreed, opting to focus on episodes where it could garner the most clinician support. In contrast, a different EI under the same AC followed the recommendation and participated in 82 episode bundles across three facilities, representing nearly 40 percent of their total Medicare volume. Other ACs were less involved in episode selection. According to our interviews, one AC provided data analytics and offered educational sessions, but ultimately allowed EIs to select their BPCI episodes.

> *"[Commercial bundle] was like a college education. We were not prepared. For BPCI, we knew what was expected and were prepared. We knew to listen to everything that our Awardee Convener had mentioned."*

In terms of the convener's role in care redesign initiatives, some participants noted that their AC held meetings for EIs to engage in discussions of best practices. One EI explained that its AC provided quality improvement experts to offer guidance on implementing care redesign. Other ACs are more actively involved in BPCI implementation onsite. During a site visit, one EI explained that its convener is entirely responsible for BPCI care redesign, data analysis, and patient management. The convener employs nursing staff at its EI hospitals. These nurses meet with the BPCI beneficiary, create a discharge plan, and maintain contact with the beneficiary throughout the episode. The EI welcomed the convener's intervention, as they did not have the resources to provide this attention on their own, and noted that PAC follow-up has proven beneficial, especially for their high-risk patient population.

d. Partners

In the context of BPCI, participants may partner with multiple types of organizations to support the initiative. During quarterly interviews and site visits we asked participants about types of partners, the role those partners play in care redesign

> *"We couldn't redesign care without [our partners] being involved. We can't do it in a vacuum."*

efforts, and the benefits and challenges of utilizing partnerships in BPCI. During the interviews, participants could define partnerships as they wished, within a broad framework that included examples such as physicians, data vendors and analysts, and other health care providers in the community that may treat BPCI patients. Interviewees described working with a variety of partners, including external data analysts, information technology (IT) consultants, local health care providers, social workers, and physicians within their own organizations.

In both the quarterly interviews and site visits, many participants described the importance of partnerships with PAC providers, especially when the PAC providers were not contractually involved in BPCI (e.g., partners not connected through formal gainsharing arrangements or through participant's organizational structure). Most participants we interviewed expressed that PAC

> *"Knowing the majority of our bundle episode cost and variations do occur within the PAC setting, any redesign success really demands key stakeholder involvement not only from those of us on the acute-care side but certainly from our PAC partners."*

partners had integral roles in implementing care redesign interventions, reducing overall costs, and improving care delivery. One participant noted that their PAC partner has successfully achieved a reduction in length of stay from a baseline of 21 days to the target of 10 days.

> *"I think the [PAC providers] have a pretty good understanding that changes are coming down the road, and like us, you either jump on the train early and help to define it, or you can continue the status quo."*

According to Awardees and EIs, the success of the acute care/PAC partnerships depends on communication and shared goals, including coordinated discharge planning, PAC provider buy-in, and the PAC partner's willingness to collaborate and change behaviors. To incentivize PAC partners, some participants use gainsharing, while others offer higher patient volumes. For example, one participant developed a CMS-approved pamphlet for patients that listed specific PAC partners. In some situations, participants have noted that PAC providers are willing to partner without additional incentives. One participant described how its PAC partners have identified the changing care delivery landscape and have been more proactive in the care coordination partnership.

Participants also described external entities and physicians as other key partners. Partnerships with external entities include contractors that provide data analytics, management, and other support activities for BPCI participants. Participants noted that these partners play significant roles in managing data systems, performing internal data analyses, facilitating data sharing across other partners, and conducting patient tracking. Partnerships with physicians also take on various forms, including gainsharing, information sharing, and standardizing care. Participants believed that positive relationships and increased engagement with physicians are critical to implementing care redesign.

Participants reported several benefits of their partnerships. For instance, participants expressed that partnerships facilitate coordination among independent entities (e.g., independent physicians, physician group practices, non-affiliated PAC providers) and allow them to "speak with one voice." Several participants also noted that partnerships enable them to focus on population health and tracking patient outcomes.

> *"While I don't think we are where we need to be yet, I think we are certainly moving in the right direction, and I think that the building of the relationships has been instrumental to that."*

Many of the participants view partnerships as a net benefit, but also report that there are challenges and limitations of the partnerships. Awardees and EIs continue to express dissatisfaction with their inability to direct beneficiaries to preferred PAC facilities. Participants have stated that their inability to control where patients receive PAC services has several downstream effects. If PACs are not receiving a significant patient volume, they are difficult to incentivize monetarily through gainsharing and are less likely to partner with Awardees and EIs to redesign care. If PACs are not invested in the care redesign efforts, it is difficult for participants to track patients after hospital

discharge. Other challenges to partnerships that participants have cited include motivating PAC providers to reduce length of stay when appropriate and coordinating data sharing across systems that may not be compatible.

Challenges with partners were also mentioned during interviews conducted with five Model 2 BPCI participants that withdrew from the initiative through June 2015. For two of these participants, challenges with partners were the key reason for exiting the program. For example, one participant, operating in a rural location, treated patients in a 300 mile radius. They found it impossible to effectively manage care for these patients after they left the participant's clinic and returned home, as patients were receiving follow-up treatment at distant facilities. The participant noted their inability to influence the care received at these facilities.

e.   *Waiver use*

The design of the BPCI initiative allows participants to use several waivers of Medicare requirements to facilitate the implementation of care redesign interventions. To use the three-day hospital stay, beneficiary incentives, or gainsharing waivers, an Awardee must describe its plans for waiver use in its Implementation Protocol (IP) to be able to utilize the waiver. An EI may or may not elect to use a waiver chosen by its Awardee. All participants have access to the telehealth and the post discharge home visit waivers. In this section we describe the use of the waivers, provide an overview of the rationales for using these waivers, and illustrate how they were implemented. Exhibit 36 provides the count of Model 2 EIs who used each of the five waivers during Q3 2014.

Exhibit 36: Participation of Episode Initiators in Various OIG/CMS Waivers, Model 2, Q3 2014

| | Model 2 EIs (N=113) | | | |
| | Model 2 EIs Allowed to Use Waiver | | Model 2 EIs that Used Waiver | |
Model 2 Waivers	**N**	**%**	**N**	**%**
Three-day hospital stay waiver	71	62.8%	29	25.7%
Beneficiary Incentives	48	42.5%	7	6.2%
Gainsharing	90	79.7%	NA*	NA*
Telehealth	NA**	NA**	0	0.0%
Home Visit	NA**	NA**	0	0.0%

*No data are available regarding use of the gainsharing waiver.
** The home visit and telehealth waivers are available to all Model 2 EIs without specifying it in their Implementation Protocols.
Sources: Lewin analysis of Awardee Implementation Protocols for Q3 2014 BPCI participants, Medicare claims data for episodes initiated Q4 2013 – Q3 2014, and Awardee-submitted data Q4 2013 – Q3 2014.
Note: Count of EIs includes both hospitals and PGPs.

Three-day hospital stay waiver

Medicare coverage rules require that beneficiaries have an inpatient hospital stay of at least three days to qualify for Medicare coverage of SNF care. With this waiver, which is only available in Model 2, hospitals may discharge beneficiaries to a SNF without meeting the three day requirement, as deemed appropriate by the treating clinicians.

> *"The [hospital three-day stay] waiver on one hand is a good thing, but on the other hand, there are some cracks in the process that can show up and then be problematic for us and our partners."*

Seventy-one EIs were allowed to use this waiver based on their Awardees' Implementation Protocols. An analysis of Medicare claims data for the BPCI episodes suggests, however, only 29 of the EIs, representing 173 episodes, used the waiver in Q3 2014.[39] Low use of the waiver has been a consistent finding, with a total of only 42 EIs, representing 443 episodes, using the waiver over the period Q4 2013 to Q3 2014.

In the quarterly interviews Awardees expressed concerns with using this waiver. Some Awardees believe that the waiver may have the unintended consequence of producing higher SNF utilization, such as when a beneficiary could instead go home with home health services. Other Awardees also expressed cautious optimism with regard to this waiver. While these Awardees believe the waiver is a positive aspect of the initiative in theory, there are concerns about the waiver's implementation, effect on partners, and potential financial liability for providers and beneficiaries if it is used incorrectly.

Beneficiary incentives

The beneficiary incentive waiver allows an EI to offer a service or product to a beneficiary that is related to the episode, but that is not typically covered by Medicare. As displayed in Exhibit 36 above, in Q3 2014, 48 EIs (roughly 43%) signed up for the beneficiary incentives waiver, allowing EIs the opportunity to provide these incentives to BPCI beneficiaries. We analyzed Awardee-submitted data that provided details about the beneficiary incentives that were distributed to beneficiaries between Q4 2013 and Q3 2014. Ten Model 2 EIs reported that they distributed incentives to 1,377 BPCI beneficiaries during this time period. These incentives ranged in value from $4.73 to $315.00 during Q4 2013 to Q3 2014. A detailed description of the distribution of beneficiary incentives by category that were provided in Q3 2014 is shown in Exhibit 37. In Q3 2014, as in the first four quarters of the initiative overall, medication management tools were most common among Model 2 Awardees that provided beneficiary incentives.

---

[39] Waiver use was identified as those BPCI episodes under Awardees allowed to use the waiver that had an anchor hospital stay of less than three days (not including the day of discharge) and a SNF stay within 30 days of hospital discharge. Episodes were excluded if they had another reason for being admitted to a SNF, such as another qualifying inpatient stay, or being in a SNF prior to the inpatient admission.

Exhibit 37: Beneficiary Incentives Waivers Distributed by EIs to Beneficiaries, Model 2, Q3 2014

Incentive Description	Awardees allowed to use incentives	EIs allowed to use incentives	Awardees that provided one or more incentive	Episodes receiving one or more incentive	Average cost per incentive provided
Transportation	10	14	3	73	$29.17
Equipment	7	7	2	94	$26.93
Home care/home visits	9	11	1	2	$283.50
Living arrangement services	3	3	0	0	NA
Telehealth/technology	6	21	0	0	NA
Wellness program/resources	6	7	0	0	NA
Medication management tools	6	6	1	158	$7.73

Source: Lewin Program Adherence Report based on analysis of Awardee-submitted data regarding disseminated beneficiary incentives in Q3 2014.

Participants continued to note that waivers could be used to address the issues in care delivery that practitioners could not affect prior to this initiative. According to those interviewed, many participants offered beneficiary incentives to ease a patient's transition to a residential care setting. For example, one Awardee described a beneficiary incentive to provide meals to patients discharged home who would otherwise be unable to cook or provide a meal on their own. Another participant noted that they covered the cost of animal boarding services if a pet presented a fall risk to the patient while they were recovering from surgery. Some Awardees, however, found implementing the beneficiary incentive waiver to be administratively challenging. These Awardees cited the requirement to track incentives distributed as a detractor to using the waiver.

Gainsharing

Gainsharing enables a participant to share any savings, with limitations, among providers with a gainsharing agreement, including its EIs. Based on a review of Awardee Implementation Protocols, 90 of the EIs active in Q3 2014 indicate an intention to participate in gainsharing and 86 of those EIs intend to gainshare with physicians.[40]

Based on the quarterly interviews, Awardees continue to view gainsharing as a tool to change practice patterns, incentivizing physicians and other partners to buy-in to the care redesign

> *"We needed the gainsharing aspect of BPCI to help align the financial incentives with the doctors to get them to the table."*

efforts. Further, Awardees have noted that gainsharing enables them to introduce quality measures that may not have been feasible without the financial incentive and to increase provider engagement in BPCI, particularly among physicians. During case studies with EIs, hospitals expressed similar sentiments. These EIs recognized that doctors play an integral role in the care redesign process, and they used gainsharing to financially reward these providers for their efforts.

---

[40] Due to a delay in collecting information from Awardees regarding gainsharing, we are unable to determine how often this waiver has been used during the first year of the initiative.

EIs also noted that gainsharing was a key element in their decision to participate in the initiative, as it enables them to offset the upfront costs of participation.

Although some participants described gainsharing with both PAC partners and individual physicians, most participants expressed that gainsharing with physicians has been more successful. Participants are able to assess physicians' baseline performance and use gainsharing to incentivize these physicians to adjust the way they provide care. EIs have also noted that gainsharing gives physicians "skin in the game," thereby encouraging them not just to improve quality metrics but to enact broader system changes. One EI suggested using gains to compensate physicians that open their schedules to patients in need of urgent care visits, while another expressed a desire to use gains to strengthen relationships with physicians that were not directly employed by the hospital.

Several participants described the challenges of gainsharing with PAC providers. One participant described how their market is saturated with SNFs and the opportunity for one of those SNFs to receive a Medicare beneficiary participating in BPCI is fairly low. Subsequently, the SNFs are not interested in

> *"Our team does not have the bandwidth to reach out to all [SNFs in the area], and they would not listen to us anyway because if they have the chance of getting one patient, they are not going to pay any attention."*

partnering with hospitals to change practice patterns for a small number of patients. Some EIs cited the SNF per diem payment schedule as a disincentive to care redesign and close coordination. EIs noted that there are other ways to engage and partner with PAC providers beyond gainsharing. For example, several EIs have incorporated preferred provider lists into their discharge planning procedures, while others have shared data and offered educational opportunities to inspire PAC facilities to change how they provide care.

> *"We particularly did not want to do gainsharing in the first year of the program… and I'll just give you my own prejudice, my experience… makes me terribly uncomfortable with gainsharing."*

Awardees that opt not to gainshare continue to cite financial risk as one of the primary inhibitors. One Awardee in particular believed that the money needed to set up gainsharing would be better invested in other services designed to support care redesign. Awardees expressed doubt as to whether gainsharing truly offered a benefit in regard to provider engagement. While some Awardees cited increased provider engagement as a reason for participating in gainsharing, others believed that they already have significant provider support and engagement, thereby making gainsharing unnecessary.

Other Awardees reported that they did not have the capacity to effectively reach out to potential gainsharing partners and implement the waiver. One Awardee described needing new full-time employees, such as a project officer, to manage the gainsharing process. EIs expressed similar concerns about the time and effort needed to establish the gainsharing process during the site visits. For some of the EIs, creating gainsharing agreements would have delayed their entrance in the BPCI initiative. Another reason cited for not using gainsharing was that Awardees want to take a wait-and-see approach, enabling them to analyze initial data and generate more interest in gainsharing from partner physicians and PAC facilities. Finally, some participants stated that gainsharing misrepresents the true goals and intentions of the BPCI initiative. These participants believe that participation in BPCI is about learning and improving care for patients, not making money.

The convener structure also impacts the decision to gainshare. Because FCs are not financially at risk and are ineligible to share in any gains, they can provide impartial guidance on gainsharing.

One FC noted that they work closely with their participants to outline a distribution methodology given the complexities of their unique market and physician network. Among the conveners that are more open to gainsharing, some have developed educational tools such as webinars where EIs can share best practices. During the site visits, EIs noted that the gainsharing methodology can significantly impact the success of the gainsharing relationship. One EI described a challenging gainsharing arrangement in which their AC retained all positive gains to cover administrative costs and care redesign services. As a result, physicians have not seen any of the financial benefits from their care interventions.

Awardees and EIs have cited the need for focused outreach and communication to ensure buy-in and participation among gainsharing entities. Further, some EIs have noted that a

> *"It's kind of new for a lot of people, so we spend a fair amount of time counseling and coaching and listening about gainsharing."*

localized approach to gainsharing may be more successful, as conveners are often outside entities with few ties to the EI and its local PAC providers. One convener has overcome this knowledge gap by analyzing data to identify high-performing PACs, leveraging relationships with national chains, and educating providers in BPCI.

Telehealth and home visit waivers

Telehealth services usually can be covered by Medicare where the originating site is one of eight health care settings located in a geographic area designated as a rural health professional shortage area or in a county not included in a Metropolitan Statistical Area. The geographic requirement is waived under the BPCI telehealth waiver, thereby making telehealth services available to beneficiaries in urban and other areas that would not usually qualify.

The post-discharge home visit waiver waives the usual requirement for direct supervision for such visits and allows participants to provide home visit services after discharge from an EI to beneficiaries who are not otherwise eligible for home health services. Home visits must be furnished by licensed clinical staff under general supervision, billed by the supervising practitioner, and provided not more than once in a 30-day episode, twice in a 60-day episode, or three times in a 90-day episode.

Because the home visit and telehealth waivers are available to all Model 2 Awardees, we do not have any information about participants' intentions to use these waivers. An analysis of Medicare claims data, however, suggests neither waiver was used for any BPCI episode in the first four quarters of the initiative.

4.   *Care redesign and cost saving strategy characteristics*

During the data submission period in Q1 2015, Model 2 EIs reported if they participated in five types of care redesign activities and the status of each intervention during 2014. The five major care redesign categories include: redesign of care pathways; enhancements in care delivery; patient activation, engagement, and risk management; care coordination; and system changes to support care. Within each care redesign category, EIs must also report how these interventions are being implemented and the progress made within the last year.

As displayed in Exhibit 38, the majority of EIs reported participating in all five types of care redesign. Participation was higher in interventions related to redesign of care pathways (95%), enhancements in care delivery (92%), patient activation, engagement, and risk management (94%),

and care coordination (94%) than participation in activities related to system changes to support care (81%). The majority of EIs participating in a care redesign activity reported that they had either begun implementation (29-44%) or their care redesign activity was fully operational (43-63%).

Exhibit 38: Participation and Status of Care Redesign Interventions, Model 2, 2014

	Redesign of Care Pathways	Enhancement in Care Delivery	Patient Activation, Engagement, and Risk Management	Care Coordination	System Changes to Support Care
Percentage of EIs that indicated participation	95%	92%	94%	94%	81%
Care redesign status (among participating EIs)					
In planning stage	9%	5%	6%	4%	10%
Implementation started	38%	37%	30%	29%	44%
Fully operational	49%	54%	60%	63%	43%
Completed & terminated	3%	4%	4%	3%	3%

Source: Lewin analysis of Awardee-submitted data collected February 2015 for 2014 Model 2 BPCI EIs.

a. Care redesign and care coordination

The case studies and quarterly interviews provided details on how participants were implementing care redesign activities across the five categories described above. Of those categories, many participants described activities related to "patient activation, engagement, and risk management."

> *"With the total joint population, it seems the more we delve into it, more and more [PAC] decisions are made by the patients well before they even see the surgeon to have the surgery. So [changing] that is a huge shift."*

Nearly all Model 2 participants at which we held a site visit mentioned the importance of patient education in their care redesign. Some interviewees indicated that participation in their educational program is mandatory for patients. There were several reasons for why this was considered important. One of the most common reasons, particularly for EIs participating in the major joint replacement episode, was that many patients have the expectation of receiving or being entitled to 21 days of care in a SNF. Because reducing PAC use was one method Model 2 participants employed to reduce the costs associated with an episode, many reported that they have increased efforts to appropriately discharge patients home with home health care rather than to a PAC facility. Some participants described patients as hesitant or resistant to the idea of being sent home directly from the hospital. Families of the patients often shared the same sentiment. Such expectations were a challenge to discharge planning.

> *"I think [BPCI] is the biggest opportunity that's come along in American health care in at least 20 years for meaningful care redesign."*

To mitigate this challenge, participants developed education programs using written materials, coaches, and classes. In addition to teaching patients how to prepare for surgery and what to expect after surgery, these programs helped to set expectations with regard to the necessary amount of PAC. Furthermore, patient education programs enabled participants to prepare their patients for successful recovery post-surgery by engaging family members, collecting information about their home environment, and helping patients to buy-in to the idea that they do not need 21 days in a SNF to have a successful recovery.

Participants also devised methods to risk-stratify their patients participating in BPCI. At several Model 2 case studies conducted over the past year, EIs indicated that they have started evaluating various factors to assess patient risk, including: psychosocial state of the home, medical history, diabetes status, presence of stairs in the home, and current medications. Using these and other factors, participants assigned patients a risk score. If a patient was designated as high risk, participants had different approaches. For some, a high-risk status resulted in changes to the discharge plan. For others, case managers only visited high-risk patients post-surgery while low-risk patients received phone call check-ups.

In both the case studies and quarterly interviews, participants frequently discussed their redesign of care pathways or clinical protocols. The most frequent change that was discussed in this context was the standardization of these protocols. Standardization referred to various themes: appropriate use of medications, standardized devices and implants, patient discharge, follow-up protocols, and pain management. Several participants also indicated that they have revised their care protocols, but these changes often predated BPCI.

> *"I think the pathway is the biggest thing that helped us standardize [the use of high cost medication] and reduce costs."*

While care standardization was a common approach to care redesign, it was also challenging, both within the hospital and among partners. Model 2 participants noted that, within a facility, care redesign was dependent on having providers that are engaged in the initiative and trained in new protocols. Many participants indicated that it was challenging to encourage providers to adopt new protocols and change the way they have practiced throughout their careers. Some sites reported implementing new care protocols for BPCI beneficiaries alone, and they noted that doing so can reduce standardization throughout the hospital at large. During case studies and quarterly interviews, Model 2 participants also identified the inability to control where their patients are treated after hospital discharge as a challenge to care standardization. Because Medicare beneficiaries maintain their choice of provider, participants indicated that they could not direct their patients to preferred PAC facilities. Thus, ensuring that PAC facilities developed and utilized new protocols was an additional challenge for Model 2 participants. Some noted that they sent hospital staff, such as patient advocates, navigators, or hospitalists, to PAC facilities to encourage the use of new care protocols, but several participants also shared that some PAC facilities were not receptive to working with hospital staff onsite.

> *"One of the major redesign features is the navigator. The orthopedics navigator position...has been instrumental in the success of the program. It really has."*

The majority of Model 2 Awardees interviewed indicated that care navigators/coordinators were a critical component of their BPCI program. In general, these individuals were responsible for meeting patients prior to surgery, discussing options for care settings after discharge, and tracking patient progress throughout the episode of care. They engaged patients in educational opportunities, ensured that discharge medications were ordered and available, and made follow-up appointments. For most Awardees, navigators and coordinators only worked with BPCI patients; some Awardees indicated that only the high-risk patients received the services of the navigator. In most cases these roles did not exist prior to BPCI. Nurses frequently filled these positions, but some Awardees mentioned that other professional staff such as therapists or social workers may also hold these roles.

Another aspect of care coordination is enhanced coordination with physicians and PAC providers. Participants indicated that they have improved the ways in which they coordinate with primary care providers, SNFs and HHAs, discharge planners, and any other staff who work with BPCI patients throughout the continuum of care. To this end, several participants implemented new IT systems that facilitated communication and data sharing. Sites operating under one convener are all making use of Pinpoint, a data collection tool that allows all providers to view and enter data on BPCI patients. In this case, sites also requested that their partner PAC facilities use this data tool to allow them to access data from across the continuum of care. Other participants noted that they increased the number of interdisciplinary team meetings so that staff would have the opportunity to discuss BPCI patients' needs and progress.

> *"One of our main approaches is to partner with the case management team and the physicians inside the hospital to determine the optimal discharge placement for the patient."*

Model 2 participants indicated that identifying BPCI-eligible patients was one of the primary challenges they faced in the initiative. For Model 2 participants, a patient was identified as a BPCI patient, and an episode was initiated, if the MS-DRG for the hospitalization was included in one of the 48 episodes included in BPCI. Until the claims data confirmed the assigned MS-DRG, however, participants indicated they were only able to assume which patients were in BPCI. Participants had to provide care based on this assumption, with the possibility that the claims data could later disqualify the patient as a bundled payment beneficiary. Those interviewed reported that patient identification was particularly difficult for patients who were admitted with one or more chronic conditions.

One of the consequences of this lag in patient identification was that it could limit the use of the three-day hospital stay waiver. One Model 2 Awardee elaborated: "If we attempt to use the three-day waiver with the CHF patient and then the coders don't code them as a CHF patient after discharge, we have put the patient in terrible financial peril."

Several participants also described difficulty determining whether a patient was already in a BPCI episode, as a patient may have previously initiated an episode at a different facility unbeknownst to the participant. Awardees cited the lack of a standard method of identifying who was in a bundled payment episode as the source of this problem.

b. Cost savings strategies

Model 2 participants indicated three general cost saving strategies: 1) device standardization; 2) reduce PAC utilization; and 3) reduce readmissions. Several respondents noted that they renegotiated device contracts because of BPCI to reduce their costs and to reduce the number of vendors. In addition, some respondents noted that their physician leadership ensured that more expensive devices were only used if clinical evidence suggested that they provided better outcomes than less expensive devices. One Awardee indicated that these changes led to a cost reduction of $8,000 per case.

Another key cost reduction strategy, managing PAC utilization, has been employed by participants since the start of the initiative. There was broad recognition among Model 2 participants that more efficient use of PAC is the best way to achieve savings under BPCI. Nearly every Model 2 participant interviewed mentioned that better managing PAC utilization – whether by reducing SNF length of stay or by appropriately substituting HHA for SNF – was part of their

plan to reduce costs. Their strategies for doing this varied. For example, participants tried to enlist the cooperation of PAC facilities to reduce the PAC length of stay. PAC provider cooperation was more likely when the participants could refer a sufficient volume of patients to the PAC providers, so they were interested in maintaining their relationships with the participants. Some participants developed "preferred provider" lists of PAC facilities that were both high-quality and supported participants' goals of reduced utilization. These lists were provided to patients and their families at discharge in order to educate them about their PAC options while also maintaining patient choice. Other participants opted to gainshare with PAC providers as an incentive to reducing length of stay.

Finally, nearly all Model 2 participants we interviewed stated that reducing readmissions was an important factor in controlling costs. This is related to the risk assessment care redesign described above, as participants noted that high-risk patients were those who were most likely to have complications or readmissions. Therefore, high-risk patients were often more closely monitored than low-risk patients. Another way that participants leveraged improved coordination with their care partners – including SNFs, physicians, and also the patient and family members – was to encourage patients to call the participant's office if they had a concern, as opposed to going straight to the emergency room.

Model 2 participants reported several challenges related to cost savings. Specifically, participants noted that bundling care for patients with chronic diseases was difficult. They explained that patients with chronic comorbidities often required more labor-intensive interventions, which often involved multiple providers. According to the Awardees we spoke with, chronic disease episodes also were less predictable and harder to manage. These challenges involved in caring for patients with chronic disease made it difficult for Awardees to reduce the length of stay and associated costs for these episodes.

Several participants also indicated that there was a large administrative burden associated with participating in BPCI. For example, several Awardees shared that a significant amount of time was required to educate staff on the goals of BPCI and to train providers in new protocols. Awardees and EIs also reported that they needed to hire more full-time employees than expected in order to meet the demands, such as increased data analysis and patient tracking, which they faced as a BPCI participant. On a related note, some participants expressed concerns that any cost savings accrued through their care redesign efforts would be insufficient to cover the overhead affiliated with hiring new staff. Awardees also indicated that BPCI-related data collection activities and analytics added to the administrative burden of their facilities.

These concerns mirror what was heard in the interviews conducted with five Model 2 BPCI participants that withdrew from the initiative as of June 2015. During these interviews, administrative burden/cost was the most frequently cited reason for terminating participation in the initiative. Several of the interviewees indicated that successful participation in BPCI required them to hire additional staff, such as data analysts, which they could not afford.

B. Impact of BPCI

This section presents the first year impact estimates of Model 2 BPCI (acute care hospital) episodes on payments, utilization, and quality of care based on episodes that were initiated during the first four quarters of the BPCI initiative (Q4 2013 through Q3 2014).[41]

We present results for the key outcomes across all Model 2 episodes in this section and then separately by clinical episode group in Sections III.D through III.K. The exhibits present the estimated differential change in risk-adjusted outcomes for patients receiving care from providers participating in BPCI between the baseline and the intervention period relative to the same change for the patients receiving care from providers in a comparison group (DiD). See Section II.D.2 for additional details on the statistical approach.

1. *Change in Medicare standardized allowed payment amounts*

We calculated the Medicare standardized allowed payment amount by type of service as well as across all services for two measurement periods. The first measurement period is the length of the episode (30-, 60-, or 90-days). We present the total payments included in the bundle and the total payments not included in the bundle by bundle length, stratified by whether or not the patient used PAC services. We also show results for total payments outcomes for the 60 days after the end of the episode (post-bundle period) and for the 30 days prior to the start of the episode (pre-bundle period.) The second measurement period is the anchor hospitalization and the 90 days after the hospital discharge (90 day PDP), regardless of the episode length. This section summarizes the results for the key payment outcomes.

Exhibit 39 presents the direction and point estimate of the DiD estimate for the key payment outcomes and each clinical episode group for Model 2. The top panel shows payment outcomes that were measured based on episodes initiated during the first year of the initiative. The bottom panel shows payment outcomes that were measured based on episodes initiated during the first three quarters of the initiative. These include post-bundle and pre-bundle payment outcomes, which were measured based on episodes initiated during the first three quarters of the initiative in order to allow for sufficient claims run out. Orange highlighted cells represent a decrease relative to the comparison group that is statistically significant at the 5 percent level. Green highlighted cells represent an increase relative to the comparison group that is statistically significant at the 5 percent level. The baseline and intervention estimates, for both BPCI and the comparison group, for key outcomes are in Sections III.D through III.K.

For most clinical episode groups, there were no statistically significant changes in standardized allowed payment between BPCI and comparison group episodes. The total standardized allowed payment for Part A and B services during the anchor hospitalization and the 90-day PDP had a statistically significant increase for patients in spinal surgery episodes ($3,477) and a statistically significant decrease for patients in orthopedic surgery episodes (-$864) treated by providers participating in BPCI relative to patients treated by the matched comparison providers. The change

---

[41] Please note that pre-bundle and post bundle period (PBP) payment outcomes and patient-assessment based outcomes are reported with a one-quarter delay and therefore are based on episodes that were initiated during the first three quarters of the BPCI initiative (Q4 2013 through Q2 2014).

in total standardized allowed payment for Part A and B services during the anchor hospitalization and the 90-day PDP was not statistically significant for the other six clinical episode groups relative to comparison providers.

While we did not observe statistically significant changes in allowed payment for most clinical episode groups, we did observe statistically significant changes in allowed payment among patients who received any PAC during the episode. There was a statistically significant decline in the total standardized allowed payment included in the bundle for patients in 30-day episodes who received PAC in three clinical episode groups -- orthopedic surgery ($-1,340), non-surgical cardiovascular ($-1,625), and cardiovascular surgery (-$4,149) -- relative to the comparison group. BPCI patients who received PAC and were in 60- or 90-day orthopedic surgery episodes also had a statistically significant decline in the total standardized allowed payment included in the bundle definition (-$2,696 and -$948, respectively).

We examined measures of the total standardized allowed payment not included in the bundle and outside of the bundle period for any indications that reduced payments were due to shifting to non-bundled services or outside of the bundle. For the clinical episode groups that experienced a statistically significant decline in total standardized allowed payment during the anchor hospitalization and the 90-day PDP or the total standardized allowed payment included in the bundle definition, there were no indications that the decline was due to shifting services outside of the bundle definition or bundle period. No clinical episode group had a statistically significant change in total amount not included in the bundle or in the post-bundle payment outcomes relative to the comparison group. There was a statistically significant increase of $50 in Part B payments during the 30 days before the start of the bundle for 90-day orthopedic surgery episodes for patients who received PAC relative to the comparison group.

Exhibit 39: Diff-in-Diff Estimate for Allowed Payment Outcomes, by Clinical Episode Group, Model 2, Baseline to Intervention

Payment Outcome	Length	PAC Use?	Orthopedic Surgery	Nonsurgical Other Medical	Nonsurgical Neurovascular	Nonsurgical Respiratory	Nonsurgical Cardiovascular	Nonsurgical & Surgical GI	Cardiovascular Surgery	Spinal Surgery
Number of episodes initiated Q4 2013 – Q3 2014			18,936	4,225	1,109	5,805	6,661	1,464	2,859	966
BPCI Cumulative Data (Q4 2013 – Q3 2014)										
Standardized allowed amount (Part A & B), IP through 90-day post-discharge period			-$864	-$96	-$194	-$32	$160	-$684	-$880	$3,477
Standardized allowed amount in bundle definition	30	No					-$956			
Standardized allowed amount in bundle definition	30	Yes	-$1,340			-$261	-$1,625		-$4,149	
Standardized allowed amount in bundle definition	60	No	$116							
Standardized allowed amount in bundle definition	60	Yes	-$2,696		$10		$533			
Standardized allowed amount in bundle definition	90	No	-$396	-$611	-$333	$98	-$389	-$48	$514	$2,025
Standardized allowed amount in bundle definition	90	Yes	-$948	-$445	$883	$194	$653	-$1,385	-$482	$2,933
Standardized allowed not included in bundle definition	90	No	-$59	$145		-$33	$38	-$96		
Standardized allowed not included in bundle definition	90	Yes	$58	$47	-$252	$125	-$143	-$227	-$140	
BPCI First Three Quarters (Q4 2013 – Q2 2014)										
Standardized allowed amount Part A & B, Days 1-30 Post bundle payment (PBP)	30	Yes	-$300				-$1,075		-$1,478	
Standardized allowed amount Part A & B, Days 1-30 PBP	90	No	-$161	$236		$169	-$140	$332	-$227	
Standardized allowed amount Part A & B, Days 1-30 PBP	90	Yes	$122	$63	-$4	$134	$59	-$395	-$482	
Standardized allowed amount Part A & B, Days 31-60 PBP	30	Yes	$167				-$1,275			
Standardized allowed amount Part A & B, Days 31-60 PBP	90	No	-$188	$351		-$137	-$197	-$31	$188	
Standardized allowed amount Part A & B, Days 31-60 PBP	90	Yes	$19	-$90	$78	-$188	-$44	-$33	-$334	
Standardized allowed amount Part B, 30-day pre-bundle payment	30	No					-$215			
Standardized allowed amount Part B, 30-day pre-bundle payment	30	Yes	-$30			-$257	-$53		$189	

Payment Outcome	Length	PAC Use?	Orthopedic Surgery	Nonsurgical Other Medical	Nonsurgical Neurovascular	Nonsurgical Respiratory	Nonsurgical Cardiovascular	Nonsurgical & Surgical GI	Cardiovascular Surgery	Spinal Surgery
Standardized allowed amount Part B, 30-day pre-bundle payment	60	No	-$747							
Standardized allowed amount Part B, 30-day pre-bundle payment	60	Yes	-$18		$292		-$54			
Standardized allowed amount Part B, 30-day pre-bundle payment	90	No	-$141	$22	-$278	-$87	$11	-$131	$67	-$207
Standardized allowed amount Part B, 30-day pre-bundle payment	90	Yes	$50	-$95	-$363	$37	$60	-$253	$141	-$103

Note: Statistical significance at the 0.05 level is indicated by orange and green shaded cells. Orange indicates the DiD estimate was negative and statistically significant; green indicates the DiD estimate was positive and statistically significant. The results presented in this table represent episodes initiated in the first four quarters of the BPCI Initiative (Q4 2013 through Q3 2014). Please note that pre-bundle and post bundle period (PBP) payment outcomes are reported with a one-quarter delay. Many of the payment outcomes are stratified by length of episode and whether or not the beneficiary had any post-acute care (PAC) use during the episode. A blank cell indicates that the outcome cannot be presented, either due to insufficient sample size or the type of episodes initiated during the time period.

Source: Lewin analysis of Medicare claims and enrollment data for episodes that began Q4 2011 through Q3 2014 for BPCI and comparison providers

We calculated a number of utilization outcomes to understand the changes that contributed to the payment outcomes summarized above. Key utilization measures include being discharged to a PAC setting and the intensity of PAC use. In particular, Exhibit 40 includes the percent of patients who were discharged to any PAC setting, the percent of those who receive PAC that were discharged to an institutional PAC setting, average inpatient length of stay (LOS), number of HH visits post anchor hospitalization discharge, and number of days during the 90 days post anchor hospitalization discharge in each institutional setting (SNF, IRF, and ACH readmissions)[42] and total number of days in any institutional setting after the anchor hospitalization discharge. Number of days of PAC use and number of HH visits are for patients who had at least one day in that setting during the 90-day post discharge period. See **Appendix L** for detailed outcome definitions.   The baseline and intervention estimates, for both BPCI and the comparison group, can be located for key outcomes in Sections III.D through III.K.

For six of the eight Model 2 clinical episode groups, the change in first PAC setting and the intensity of PAC use for patients treated by BPCI EIs was not statistically significant relative to the change for the patients treated by comparison providers. Among patients treated by BPCI EIs with orthopedic surgery episodes, there was a statistically significant decrease in the proportion of patients receiving PAC who were discharged to an institutional PAC setting relative to the comparison group (a decline of 4.9 percentage points). In addition, among those who were discharged to a SNF, there was a statistically significant decline of 1.3 days in the number of days in the SNF relative to the comparison group. The average LOS for the acute inpatient stay also had a statistically significant decrease of 0.1 days relative to the comparison group.

Patients treated by BPCI EIs with cardiovascular surgery episodes also experienced a statistically significant change in first PAC setting relative to the comparison group. The proportion of patients treated by BPCI EIs with cardiovascular surgery episodes with any PAC use discharged to institutional PAC had a statistically significant decline (9.9 percentage points) relative to patients in the comparison group. In addition, there was a statistically significant increase in HH visits (1.5 visits) for patients who were discharged to HHA relative to the comparison group. These statistically significant shifts in PAC use did not, however, result in a statistically significant decline in total standardized allowed payment for Part A and B services during the anchor hospitalization and the 90-day PDP.

---

[42] There was insufficient sample size to examine average LTCH length of stay.

Exhibit 40: Diff-in-Diff Estimate for Utilization Outcomes, by Clinical Episode Group, Model 2, Baseline to Intervention

Measure	Orthopedic Surgery	Nonsurgical Other Medical	Nonsurgical Neurovascular	Nonsurgical Respiratory	Nonsurgical Cardiovascular	Nonsurgical & Surgical GI	Cardiovascular Surgery	Spinal Surgery
Number of episodes initiated Q4 2013 – Q3 2014	18,936	4,225	1,109	5,805	6,661	1,464	2,859	966
% discharged to post-acute care	0.1 pp	0.5 pp	-1.8 pp	0.2 pp	1.4 pp	1.1 pp	-0.1 pp	0.6 pp
% discharged to an institution out of those who received any post-acute care	-4.9 pp	1.9 pp	0.8 pp	1.9 pp	-0.4 pp	1.1 pp	-9.9 pp	-9.2 pp
Acute Inpatient Care LOS	-0.1	0.1	0.0	0.0	0.0	0.0	0.0	0.2
Number of HH Visits, 90-day PDP[1]	0.1	0.5	1.6	1.2	0.6	0.0	1.5	0.4
Number of Institutional Days, 90-day PDP[1]	-0.5	-0.6	1.7	0.1	0.1	-0.8	-1.1	3.4
Number of IRF Days, 90-day PDP[1]	0.2		-0.3	1.7	-0.1		1.6	1.7
Number of Readmission Days, 90-day PDP[1]	0.1	-0.3	-0.1	0.2	0.1	-1.1	0.3	1.4
Number of SNF Days, 90-day PDP[1]	-1.3	-1.1	4.2	-0.2	0.0	1.1	-0.3	2.3

[1] The number of visits/days is conditional on having at least 1 day in the respective setting.

Note: Statistical significance at the 0.05 is indicated by orange and green shaded cells. Orange indicates the DiD estimate was negative and statistically significant; green indicates the DiD estimate was positive and statistically significant. The results presented in this table represent episodes initiated in the first four quarters of the BPCI Initiative (Q4 2013 through Q3 2014). A blank cell indicates that the outcome cannot be presented, either due to insufficient sample size or the type of episodes initiated during the time period. pp=percentage points.

Source: Lewin analysis of Medicare claims and enrollment data for episodes that began Q4 2011 through Q3 2014 for BPCI and comparison providers.

2. *Changes in beneficiary quality of care*

Exhibit 41 presents the impact of Model 2 BPCI on key claims-based and patient assessment-based quality of care measures by clinical episode group. The outcomes are stratified in the exhibit based on the quarters included in the DiD estimate. Claim-based outcomes were measured based on episodes initiated during the first year of the initiative while assessment-based outcomes were measured based on episodes initiated during the first three quarters of the initiative in order to allow for sufficient assessment data run out. Orange highlighted cells represent a statistically significant decrease relative to the comparison group. Green highlighted cells represent a statistically significant increase relative to the comparison group.

The claims-based quality of care measures are mortality within 30 days of discharge and emergency department use and readmission rates within 30 days and 90 days of discharge from the anchor hospitalization. The patient assessment outcomes reflect change in functional status among beneficiaries who received post-acute care. See **Appendix L** for detailed outcome definitions. The baseline and intervention estimates, for both BPCI and the comparison group, can be located for key outcomes in Sections III.D through III.K.

a. *Claim-based measures*

There were no statistically significant differences in the change in mortality, readmission rates, or emergency department visits between beneficiaries treated by BPCI providers and beneficiaries treated by comparison providers, with three exceptions. For beneficiaries in the cardiovascular surgery clinical episodes mortality rate increased from 1.6% to 1.9% for BPCI beneficiaries while it decreased for beneficiaries treated by comparison providers (2.1% vs. 1.4%), although this result did not persist with additional data and a more refined comparison group.[43] The emergency department visit rate during the 30-day PDP increased from 9.3% to 12.0% for BPCI beneficiaries in cardiovascular surgery episodes, compared with an increase of 10.8 to 11.3% for comparison group beneficiaries, although the change in the emergency department visit rate during the 90-day PDP did not differ significantly from the comparison group. The mortality rate for BPCI beneficiaries with spinal surgery episodes declined from 1.2 % to 0.2%, compared with an increase from 0.5% to 1.2% for beneficiaries treated by comparison providers, although this result did not appear to persist with additional data and a more refined comparison group.[44]

b. *Assessment-based measures*

The majority of the assessment-based quality measures did not change in a statistically significant way among the BPCI Model 2 beneficiaries, relative to the comparison beneficiaries.

---

[43] Given the importance of this finding, it was reevaluated with an additional nine months of data after the comparison group methodology had been refined. The results based on Q4 2013 to Q2 2015 data and the refined comparison group indicate that for three of the component episodes in the cardiovascular surgery clinical episode group, BPCI beneficiaries did not have a statistically significant change (at the 5%, 10%, or 20% levels) in the 30-day mortality rate relative to comparison beneficiaries.

[44] The 90-day mortality rates from Q4 2013 to Q2 2015 for BPCI beneficiaries with spinal fusion episodes (a component of the spinal surgery clinical episode group) did not have a statistically significant change (at the 5%, 10%, or 20% levels) relative to comparison beneficiaries. The 30-day mortality rate was not available because there were too few mortality events to produce risk-adjusted results.

Among beneficiaries whose first PAC setting was a SNF or an IRF, we did not observe any statistically significant changes in assessment-based measures among beneficiaries treated by BPCI providers, relative to the comparison group.

For those whose first PAC provider was a HHA, we observed two statistically significant changes relative to the comparison group. First, among patients with orthopedic surgery episodes, the BPCI intervention was associated with a statistically significant decrease of 1.6 percentage points in the share of patients that exhibited improvement in upper body dressing, relative to the comparison group. This was caused by a decline of 0.5 percentage points in share of patients that exhibited improvement in upper body dressing from baseline to intervention period among BPCI patients combined with an increase of 1.1 percentage points in the same rate among comparison patients. Second, among patients with non-surgical and surgical GI episodes, BPCI providers had a statistically significant decrease of 12.0 and 14.3 percentage points in the share of patients that exhibited improvement in upper body dressing and lower body dressing, respectively, relative to the comparison group. Both statistically significant program effects were caused by a decrease in rates among BPCI patients and an increase in rates among comparison patients.

Exhibit 41: Diff-in-Diff Estimate for Claim-based and Assessment-based Quality Outcomes, by Clinical Episode Group, Model 2, Baseline to Intervention

Measure	Orthopedic Surgery	Nonsurgical Other Medical	Nonsurgical Neurovascular	Nonsurgical Respiratory	Nonsurgical Cardiovascular	Nonsurgical & Surgical GI	Cardiovascular Surgery	Spinal Surgery
Number of episodes initiated Q4 2013 – Q3 2014	18,936	4,225	1,109	5,805	6,661	1,464	2,859	966
BPCI Cumulative to Date (Q4 2013 – Q3 2014)								
All-cause mortality rate, 30-day PDP	-0.0 pp	-1.5 pp	-1.7 pp	-0.4 pp	-0.5 pp	0.3 pp	0.9 pp	-1.6 pp
Emergency Department Use, 30-day PDP	0.2 pp	0.4 pp	0.7 pp	-0.1 pp	-0.4 pp	1.4 pp	2.2 pp	2.5 pp
Emergency Department Use, 90-day PDP	0.1 pp	1.6 pp	0.1 pp	0.5 pp	-0.2 pp	0.1 pp	0.4 pp	2.2 pp
Unplanned readmission rate, 30 day PDP	0.0 pp	0.7 pp	0.5 pp	0.3 pp	0.5 pp	-0.8 pp	-1.3 pp	0.7 pp
Unplanned readmission rate, 90-day PDP	-0.1 pp	0.5 pp	-1.0 pp	0.1 pp	1.3 pp	1.4 pp	-0.7 pp	1.6 pp
BPCI First Three Quarters (Q4 2013 – Q2 2014)								
ADL HHA, improved ambulation	-1.6 pp	-3.0 pp	7.4 pp	3.9 pp	3.2 pp	-0.7 pp	3.7 pp	3.8 pp
ADL, HHA, improved bathing	-0.5 pp	-4.0 pp	-12.0 pp	0.7 pp	-2.8 pp	-11.7 pp	-3.2 pp	0.4 pp
ADL, HHA, improved bed transferring	-0.9 pp	4.7 pp	5.6 pp	2.9 pp	5.0 pp	-9.7 pp	1.4 pp	-1.7 pp
ADL, HHA, improved lower body dressing	-1.2 pp	0.8 pp	-5.0 pp	3.4 pp	-1.3 pp	-14.3 pp	-3.8 pp	-7.1 pp
ADL, HHA, improved upper body dressing	-1.6 pp	-5.0 pp	0.6 pp	1.3 pp	1.3 pp	-12.0 pp	-2.1 pp	4.2 pp
ADL, IRF, average change in mobility score	0.0 pp	0.1 pp	0.1 pp	-0.5 pp	0.2 pp		-0.4 pp	0.1 pp
ADL, IRF, average change in self-care score	0.5 pp	-1.5 pp	-0.1 pp	-2.1 pp	0.4 pp		0.2 pp	-0.8 pp
ADL, SNF, improved mobility function	-3.5 pp	1.3 pp	7.9 pp	0.7 pp	-1.8 pp	-0.2 pp	9.1 pp	10.9 pp
ADL, SNF, improved overall function	-2.2 pp	3.2 pp	0.8 pp	2.5 pp	-0.4 pp	10.5 pp	2.8 pp	5.1 pp
ADL, SNF, improved self-care function	-2.7 pp	-3.3 pp	4.8 pp	0.3 pp	2.8 pp	13.0 pp	6.0 pp	-0.2 pp

Note: Statistical significance at the 0.05 level is indicated by orange and green shaded cells. Orange indicates the DiD estimate was negative and statistically significant; green indicates the DiD estimate was positive and statistically significant. The results presented in this table represent episodes initiated in the first four quarters of the BPCI Initiative (Q4 2013 through Q3 2014). Please note that assessment-based quality measures are reported with a one-quarter delay. A blank cell indicates that the outcome cannot be presented, either due to insufficient sample size or the type of episodes initiated during the time period. pp=percentage points.

Source: Lewin analysis of Medicare claims and enrollment data for episodes that began Q4 2011 through Q3 2014 and assessment data for episodes that began Q4 2011 through Q2 2014 for BPCI and comparison providers.

c. Beneficiary survey

Exhibits 42 and 43 below report the results for self-reported functional status improvement, functional status decline, and patient care experiences, respectively for all Model 2 beneficiary survey strata. Response rates for all Model 2 strata can be found in **Appendix O.**

Overall, there were few statistically significant differences in functional status, other than walking without rest and use of stairs, between BPCI Model 2 respondents and comparison respondents. . Overall, BPCI respondents were more likely than comparison respondents to report improvement in using stairs: a difference of 6.0 percentage points, statistically significant at 0.05. This difference was primarily driven by respondents without MCC, among whom the difference was 8.3 percentage points. Relative to comparison respondents, BPCI respondents with a MJRLE episode were 8.2 percentage points more likely to report improvement in walking without rest and 7.5 percentage points more likely to report improvement in the use of stairs. Both estimates were statistically significant.

BPCI respondents in the non-MCC stratum were also less likely than comparison respondents to decline in their ability to use stairs, a difference of 7.4 percentage points, statistically significant at 0.05.. BPCI respondents with a non-surgical respiratory episode were 7.1 percentage points less likely to decline in their ability to walk without rest. There were no other statistically significant differences in the rates of improvement or decline for any of the other functional status measures.

Exhibit 42: Improvement and Decline in Functional Status for BPCI and Comparison Survey Respondents, Model 2, June, October, and November 2014

Functional Measure	Survey Strata[a]	BPCI Rate (n)	Comparison Rate (n)	Treatment Effect
Improvement in bathing, dressing, using the toilet, or eating	Overall	72.6% (712)	73.2% (708)	-0.5 [-4.8, 3.7]
	MCC	59.7% (278)	60.5% (296)	-0.9 [-8.1, 6.4]
	Non-MCC	81.9% (434)	82.4% (412)	-0.5 [-5.6, 4.6]
	MJRLE	85.0% (353)	84.5% (373)	0.5 [-4.3, 5.2]
	Cardio	60.7% (347)	57.2% (342)	3.5 [-2.5, 9.4]
	Respiratory	60.3% (266)	59.7% (271)	0.6 [-6.2, 7.4]
Improvement in walking without rest	Overall	82.0% (711)	80.3% (721)	1.7 [-2.1, 5.4]
	MCC	73.0% (278)	70.8% (303)	2.2 [-4.4, 8.8]
	Non-MCC	88.3% (433)	87.6% (418)	0.7 [-3.4, 4.9]
	MJRLE	65.7% (351)	57.5% (374)	8.2* [1.6, 14.8]
	Cardio	27.4% (342)	27.2% (347)	0.2 [-5.6, 6.0]
	Respiratory	30.8% (262)	29.2% (269)	1.6 [-5.1, 8.2]

Functional Measure	Survey Strata[a]	BPCI Rate (n)	Comparison Rate (n)	Treatment Effect
Improvement in use of mobility device (i.e., less frequent)	Overall	83.3% (714)	85.2% (716)	-1.9 [-5.6, 1.7]
	MCC	74.0% (280)	77.8% (297)	-3.8 [-10.1, 2.4]
	Non-MCC	90.2% (434)	90.4% (419)	-0.2 [-4.4, 4.0]
	MJRLE	60.8% (354)	63.5% (373)	-2.7 [-9.2, 3.8]
	Cardio	37.7% (345)	40.2% (349)	-2.5 [-7.7, 2.8]
	Respiratory	41.7% (268)	44.1% (270)	-2.5 [-7.4, 2.4]
Improvement in using stairs	Overall	61.3% (717)	55.3% (718)	6.0* [1.8, 10.2]
	MCC	46.0% (282)	42.4% (302)	3.6 [-3.1, 10.4]
	Non-MCC	72.7% (435)	64.4% (416)	8.3* [3.0, 13.5]
	MJRLE	65.4% (353)	57.9% (366)	7.5* [0.9, 14.1]
	Cardio	26.8% (335)	29.4% (333)	-2.5 [-8.5, 3.4]
	Respiratory	28.2% (261)	24.4% (265)	3.9 [-3.2, 11.0]
Improvement in planning regular tasks	Overall	63.5% (721)	67.2% (711)	-3.7 [-8.0, 0.7]
	MCC	50.0% (284)	54.4% (295)	-4.4 [-11.6, 2.8]
	Non-MCC	73.3% (437)	76.4% (416)	-3.2 [-8.5, 2.2]
	MJRLE	76.0% (355)	80.0% (377)	-4.0 [-9.4, 1.4]
	Cardio	46.4% (345)	47.7% (353)	-1.3 [-7.4, 4.8]
	Respiratory	44.8% (267)	44.4% (274)	0.5 [-6.8, 7.7]
Improvement in physical/emotional problems limiting social activities (i.e., less frequent)[b]	MJRLE	75.8% (348)	72.0% (372)	3.8 [-2.4, 10.0]
	Cardio	45.3% (337)	51.0% (342)	-5.6 [-13.0, 1.6]
	Respiratory	45.1% (271)	45.8% (273)	-0.7 [-8.6, 7.2]
Improvement in pain limiting regular activities (i.e., less frequent)[b]	MJRLE	81.9% (355)	77.9% (378)	4.0 [-1.6, 9.7]
	Cardio	45.4% (338)	46.6% (349)	-1.2 [-8.0, 5.6]
	Respiratory	44.0% (266)	42.7% (272)	1.3 [-6.6, 9.2]

Functional Measure	Survey Strata[a]	BPCI Rate (n)	Comparison Rate (n)	Treatment Effect
Decline in bathing, dressing, using the toilet, or eating	Overall	16.0% (712)	15.1% (708)	0.9 [-2.8, 4.6]
	MCC	24.5% (278)	25.5% (296)	-1.0 [-8.0, 6.1]
	Non-MCC	9.7% (434)	7.7% (412)	2.0 [-2.3, 6.2]
	MJRLE	8.2% (353)	8.4% (373)	-0.1 [-3.9, 3.7]
	Cardio	19.8% (344)	22.3% (340)	-2.5 [-8.6, 3.6]
	Respiratory	18.1% (266)	21.1% (271)	-3.0 [-9.4, 3.4]
Decline in walking without rest	Overall	8.2% (643)	8.5% (657)	-0.3 [-3.5, 2.9]
	MCC	12.4% (238)	14.4% (260)	-2.0 [-8.2, 4.2]
	Non-MCC	5.5% (390)	5.1% (381)	0.4 [-2.8, 3.6]
	MJRLE	14.4% (351)	15.7% (374)	-1.3 [-6.2, 3.5]
	Cardio	37.4% (342)	37.0% (347)	0.4 [-5.6, 6.3]
	Respiratory	27.9% (262)	34.9% (269)	-7.1* [-13.5, -0.6]
Decline in use of mobility device (i.e., more frequent)	Overall	12.1% (714)	11.0% (716)	1.1 [-2.2, 4.5]
	MCC	19.0% (280)	15.6% (297)	3.4 [-2.5, 9.2]
	Non-MCC	7.0% (434)	7.7% (419)	-0.6 [-4.5, 3.2]
	MJRLE	26.7% (354)	24.2% (373)	2.5 [-3.3, 8.3]
	Cardio	48.3% (345)	44.5% (349)	3.8 [-2.1,9.7]
	Respiratory	42.5% (268)	40.1% (270)	2.4 [-3.5, 8.2]
Decline in using stairs	Overall	31.5% (717)	37.2% (718)	-5.6* [-9.6, -1.7]
	MCC	46.7% (282)	50.7% (302)	-4.0 [-10.7, 2.8]
	Non-MCC	20.2% (435)	27.6% (416)	-7.4* [-12.0, -2.9]
	MJRLE	15.4% (353)	16.7% (366)	-1.3 [-6.0, 3.4]
	Cardio	42.0% (335)	43.7% (333)	-1.7 [-7.8, 4.5]
	Respiratory	38.5% (261)	42.9% (265)	-4.4 [-10.4, 1.6]

Functional Measure	Survey Strata[a]	BPCI Rate (n)	Comparison Rate (n)	Treatment Effect
Decline in planning regular tasks	Overall	14.8% (632)	12.5% (631)	2.4 [-1.6, 6.4]
	MCC	20.5% (230)	19.3% (243)	1.2 [-6.4, 8.8]
	Non-MCC	11.2% (397)	8.4% (383)	2.8 [-1.5, 7.1]
	MJRLE	13.7% (355)	9.5% (377)	4.2 [-0.1, 8.5]
	Cardio	27.5% (342)	30.1% (351)	-2.7 [-9.0, 3.7]
	Respiratory	25.6% (267)	25.7% (274)	-0.1 [-6.6, 6.3]
Decline in physical/emotional problems limiting social activities (i.e., more frequent)[b]	MJRLE	10.1% (348)	12.4% (372)	-2.3 [-6.0, 3.4]
	Cardio	27.8% (330)	27.5% (334)	0.3 [-6.4, 7.0]
	Respiratory	27.9% (271)	30.5% (273)	-2.5 [-9.8, 4.7]
Decline in pain limiting regular activities (i.e., more frequent)[b]	MJRLE	7.0% (355)	8.5% (378)	-1.5 [-5.4, 2.5]
	Cardio	20.8% (338)	21.7% (349)	-0.9 [-6.8, 5.0]
	Respiratory	242% (266)	24.6% (272)	-0.5 [-7.2, 6.3]

* Denotes statistical significance at the 5% level.

Notes: Treatment effect reported in percentage points. 95% confidence interval is in brackets. Survey results were weighted to adjust for sampling and nonresponse.

MJRLE refers to "major joint replacement of the lower extremity," *cardio* refers to "non-surgical cardiovascular", and *respiratory* refers to "non-surgical respiratory" episodes.

MCC refers to episodes with major complicating conditions; non-MCC refers to episodes without a major complicating condition.

The questions and possible responses for walking, use of stairs, and use of a mobility device changed drastically between Waves 1 and 2. Therefore, rates of improvement and decline for these measures in the overall, MCC, and non-MCC strata (from Wave 1) may differ substantially from those in the MJRLE, Cardio, and Respiratory strata (from Wave 2).

BPCI and comparison samples sizes for each question are reported in parentheses. The sample sizes vary between questions due to varying rates of item-response.

[a] The overall, MCC, and non-MCC strata were sampled in June 2014. The MJRLE, Cardio, and Respiratory strata were sampled in October and November 2014.

[b] Rates of change in these measures can only be estimated for the October/November 2014 sample.

Source: Lewin analysis of Medicare beneficiary survey for BPCI and comparison group survey respondents.

In general, there were few statistically significant differences in health care experience outcomes between BPCI Model 2 respondents and comparison respondents. Results indicate that some self-reported health outcomes may be improving among BPCI respondents, but certain health care experience measures may be deteriorating. However, there was no statistically significant difference in overall satisfaction with recovery between BPCI and comparison respondents.

BPCI respondents in the non-MCC stratum were 3.8 percentage points more likely to agree that medical staff clearly explained to them how to take their medications before they went home. BCPI respondents with a MJRLE episode were 2.8 percentage points less likely than comparison respondents to agree or strongly agree that medical staff clearly explained what follow-up appointments or treatments would be needed. In addition, BPCI respondents with a non-surgical cardiovascular episode were 4.5 percentage points less likely than similar comparison respondents to agree or strongly agree that they had been able to manage their own health needs since discharge. We note that while these differences were statistically significant, affirmative responses for each question were over 90% in each group, suggesting that overall satisfaction with care experience remained high.

Exhibit 43: Health Care Experience for BPCI and Comparison Survey Respondents, Model 2, June, October, and November 2014

Health Care Experience Measures	Survey Strata[a]	BPCI Rate (n)	Comparison Rate (n)	Treatment Effect
Thinking about all the care you received in the hospital before and afterwards, how often did you, your family, or your caregiver get conflicting advice from medical staff about your treatment? 1[response = Never]	Overall	75.5% (730)	71.2% (724)	4.3 [-0.2, 8.8]
	MCC	69.4% (291)	63.4% (304)	6.1 [-1.7, 13.8]
	non-MCC	80.0% (439)	77.1% (420)	2.8 [-2.6, 8.2]
	MJRLE	81.2% (357)	83.0% (379)	-1.9 [-7.5, 3.8]
	Cardio	65.5% (341)	70.3% (352)	-4.8 [-11.9, 2.3]
	Respiratory	65.3% (274)	69.0% (274)	-3.7 [-11.8, 4.4]
Thinking about all of the care you received in the hospital and afterwards, how often were the services you got appropriate for the level of care you needed? 1[response = Always]	Overall	65.5% (717)	61.8% (713)	3.7 [-1.3, 8.8]
	MCC	58.7% (285)	55.2% (302)	3.5 [-4.7, 11.7]
	non-MCC	70.5% (432)	66.9% (411)	3.7 [-2.7, 10.0]
	MJRLE	67.6% (358)	70.1% (378)	-2.4 [-9.2, 4.4]
	Cardio	55.9% (343)	63.1% (356)	-7.2 [-14.6, 0.2]
	Respiratory	55.6% (273)	58.5% (271)	-2.9 [-11.5, 5.7]

Health Care Experience Measures	Survey Strata[a]	BPCI Rate (n)	Comparison Rate (n)	Treatment Effect
Thinking about when you left the hospital, were you discharged at the right time? 1[response = Yes]	Overall	89.5% (719)	88.3% (715)	1.2 [-2.2, 4.6]
	MCC	87.2% (284)	86.0% (299)	1.2 [-4.6, 7.0]
	non-MCC	91.2% (435)	90.0% (416)	1.1 [-2.9, 5.2]
	MJRLE	89.9% (358)	93.3% (377)	-3.4 [-7.4, 0.7]
	Cardio	86.3% (348)	86.0% (355)	0.3 [-5.0, 5.5]
	Respiratory	86.7% (274)	83.0% (272)	3.7 [-2.8, 10.1]
Thinking about when you left the hospital listed in the cover letter, the medical staff took your preferences and those of your family or your caregiver into account in deciding what health care services you should have after you left the hospital. 1[response = Agree/Strongly Agree]	Overall	93.5% (640)	93.2% (610)	0.2 [-2.5, 3.0]
	MCC	95.4% (251)	92.9% (257)	2.4 [-1.7, 6.6]
	non-MCC	92.1% (389)	93.5% (353)	-1.3 [-5.0, 2.3]
	MJRLE	95.9% (334)	95.0% (354)	0.8 [-2.4, 4.0]
	Cardio	92.4% (307)	95.7% (308)	-3.7 [-7.3, 0.5]
	Respiratory	92.1% (226)	90.8% (236)	1.3 [-3.7, 6.2]
Before you prepared to go home (or to someone else's home, or to an assisted living facility), you and your family or caregiver had a good understanding of how to take care of yourself. 1[response = Agree/Strongly Agree]	Overall	94.7% (649)	95.0% (640)	-0.2 [-2.7, 2.2]
	MCC	93.8% (235)	95.4% (257)	-1.6 [-5.7, 2.4]
	non-MCC	95.3% (414)	94.7% (383)	0.7 [-2.4, 3.7]
	MJRLE	96.5% (342)	96.6% (357)	-0.2 [-2.9, 2.6]
	Cardio	95.2% (297)	96.9% (315)	-1.7 [-5.0, 1.6]
	Respiratory	95.3% (238)	94.7% (227)	0.6 [-3.6, 4.8]
Before you prepared to go home (or to someone else's home, or to an assisted living facility), medical staff clearly explained how to take your medications. 1[response = Agree/Strongly Agree]	Overall	95.4% (644)	94.0% (616)	1.4 [-1.1, 3.9]
	MCC	92.8% (240)	94.8% (254)	-2.0 [-6.3, 2.3]
	non-MCC	97.2% (404)	93.4% (362)	3.8* [0.7, 6.8]
	MJRLE	95.1% (330)	96.0% (350)	-1.0 [-4.1, 2.2]
	Cardio	95.8% (297)	93.7% (306)	2.1 [-1.8, 5.9]
	Respiratory	96.5% (236)	93.9% (222)	2.7 [-1.3, 6.6]

Health Care Experience Measures	Survey Strata[a]	BPCI Rate (n)	Comparison Rate (n)	Treatment Effect
Before you prepared to go home (or to someone else's home, or to an assisted living facility), medical staff clearly explained what follow-up appointments or treatments would be needed 1[response = Agree/Strongly Agree]	Overall	96.4% (663)	95.0% (639)	1.4 [-0.9, 3.6]
	MCC	94.8% (246)	94.5% (258)	0.3 [-3.6, 4.3]
	non-MCC	97.4% (417)	95.4% (381)	2.0% [-0.7, 4.8]
	MJRLE	96.9% (332)	99.7% (341)	-2.8* [-4.7, -0.8]
	Cardio	93.7% (293)	95.5% (310)	-1.8 [-5.5, 2.0]
	Respiratory	95.4% (225)	97.4% (226)	-2.0 [-5.4, 1.3]
Overall, since you returned home (or to someone else's home, or to an assisted living facility), you and your caregivers have been able to manage your health needs. 1[response = Agree/Strongly Agree]	Overall	96.0% (660)	96.2% (648)	-0.2 [-2.3, 1.9]
	MCC	95.4% (242)	94.6% (265)	0.8 [-2.9, 4.6]
	non-MCC	96.4% (418)	97.4% (383)	-1.0 [-3.5, 1.5]
	MJRLE	99.7% (331)	98.6% (352)	1.0 [-0.3, 2.4]
	Cardio	92.2% (300)	96.7% (311)	-4.5* [-8.4, -0.6]
	Respiratory	92.8% (230)	94.6% (230)	-1.8 [-6.7, 3.0]
Overall, how satisfied are you with your recovery since you left the hospital? 1[response = Extremely/Quite a Bit]	Overall	69.9% (719)	69.6% (699)	0.3 [-4.7, 5.2]
	MCC	61.4% (285)	64.7% (291)	-3.3 [-11.4, 4.9]
	non-MCC	76.0% (434)	73.3% (408)	2.7 [-3.4, 8.8]
	MJRLE	82.8% (341)	82.9% (363)	-0.1 [-5.8, 5.6]
	Cardio	61.4% (326)	66.7% (332)	-5.4 [-12.9, 2.1]
	Respiratory	61.9% (259)	63.6% (257)	-1.8 [-10.2, 6.7]

* Denotes statistical significance at the 5% level.

Notes: Results are unadjusted. Treatment effect reported in percentage points. 95% confidence interval is in brackets. Survey results were weighted to adjust for sampling and nonresponse.

MJRLE refers to "major joint replacement of the lower extremity," *cardio* refers to "non-surgical cardiovascular", and *respiratory* refers to "non-surgical respiratory" episodes.

BPCI and comparison samples sizes for each question are reported in parentheses. The sample sizes vary between questions due to varying rates of item-response.

[a] The overall, MCC, and non-MCC strata were sampled in June 2014. The MJRLE, Cardio, and Respiratory strata were sampled in October and November 2014.

Source: Lewin analysis of Medicare beneficiary survey for BPCI and comparison group survey respondents.

C. Market Dynamics

We created two hospital-to-PAC provider concentration indices to determine whether BPCI beneficiaries were discharged to fewer PAC providers after discharge from a hospital EI. These indexes are analogous to a Herfindahl Index but applied to PAC care for BPCI-patients discharged from a Model 2 EI hospital (see section II.D.5.d.). We also calculated the market share of Model 2 EIs to determine whether EIs have captured a greater share of BPCI-eligible episodes over time. We hypothesized that Model 2 EIs would leverage existing relationships with PAC providers and focus on encouraging their patients towards these providers in order to better manage their patients' care across multiple settings. We cannot determine whether changes in the concentration indices were due to changes in BPCI providers' activities alone, since BPCI patients retain their freedom to choose their PAC provider.

For the hospital-to-SNFs concentration index, we calculated the proportions of a hospital EI's patients whose first PAC stay was at each SNF in the market, squared the ratios and summed them across all SNFs. This index can range from zero to one, with a higher number indicating greater concentration, that is, the EI's patients discharged to fewer SNFs. A hospital-to-HHAs concentration index was constructed in the same manner.

A third measure, the market share of hospital EIs, was calculated as the number of BPCI-eligible episodes admitted to a hospital EI, divided by the total number of the same type of episodes admitted to all other hospitals in the market.   We hypothesized that the hospital EIs would increase their market share of BPCI-eligible episodes. As discussed in the Methods section, we defined the BPCI market as the CBSA in which it is located.

We calculated these measures separately for three clinical episode groups, since market share and PAC patterns can be considerably different for different clinical conditions. The three clinical episode groups are MJRLE, CHF, and sepsis.   The sections below contain results for MJRLE episodes; results for CHF and sepsis are in **Appendix R**. We focus on MJRLE because it is the largest volume clinical episode group under all BPCI Models and offers the best opportunity to observe changes that may be due to BPCI.

The two subsections below present time-trend analyses using two years of pre-BPCI data and one year of data after BPCI was implemented, segmented into 6-month intervals. The first subsection presents descriptive statistics for the MJRLE hospital-to-SNFs concentration index and hospital-to-HHAs concentration index. The second sub-section presents statistics for the market share measure. Within each table we present national results and also for selected individual markets. For Model 2 hospital EIs we explored the following individual markets/CBSAs: New York-Wane-White Plains, NY-NJ (New York market hereafter), Phoenix-Mesa-Scottsdale, AZ (Phoenix market), and Providence-New Bedford-Fall River, RI-MA (Providence market).

This is an exploratory analysis based on data from the first cohorts of BPCI EIs and the first year of the BPCI program. These results will be used to refine our measures and approach. In future reports we will expand this analysis to reflect additional EIs that joined the program, greater volume of BPCI episodes, and more time for EIs to implement strategic changes in response to BPCI incentives.

1. *Hospital-to-SNFs concentration index for Model 2 EIs*

Exhibits 44 and 45 present the hospital-to-SNFs and hospital-to-HHAs concentration indices trends for MJRLE episodes over all Model 2 EIs. These tables also include the average highest percent of admissions from each hospital EI at one PAC provider and the average number of PAC providers to which each EI's patients were admitted. For context, we also report the total number of SNF or HHA providers with MJRLE admissions from BPCI hospital EIs in each individual market.[45] **Appendix P** shows the same results for CHF and sepsis episodes.

There was no indication of increasing concentration. The average hospital-to-SNFs concentration index for MJRLE patients was 0.27 in Q2 2014/Q3 2014, little different from that during the pre-BPCI periods.[46] The median index values over all EIs were lower than the mean index values and also show no BPCI effect on the nationwide trend. We did not observe any meaningful change in the average number of SNF providers each EI's patients went to, or in the average highest percent of patients admitted to one SNF.

The same was true for MJRLE patients discharged from Model 2 hospital EIs and admitted to HHAs. The average hospital-to-HHAs concentration index in Q2 2014/Q3 2014 was 0.46 for MJRLE, which was the same as the pre-BPCI rate. There was little change in the average concentration index, median concentration index, average number of HHA providers to which each EI's patients were admitted, or average highest percent of patients admitted to one HHA.

We examined trends in three individual markets of different sizes and composition, which each had a relatively large number of BPCI EIs. Again, we saw no change in the concentration of MJRLE patients admitted to SNFs or HHAs after being discharged from EI hospitals. We did not observe any substantive differences between the pre-BPCI and one year intervention periods in the average number of SNF or HHAs admitting these MJRLE patients, or the average highest percent of patients admitted to one SNF or HHA.

Appendix P presents similar findings at both the national and the individual market levels for CHF and sepsis patients. We did see some indication of a declining hospital-to-SNFs concentration index for CHF patients and for the hospital-to-HHAs index for sepsis patients. Both of these average concentration indexes were lower in the Q4 2013/Q1 2014 and Q2 2013 / Q3 2014 periods than in any pre-BPCI period. At the all-market level, the average hospital-to-SNFs and hospital-to-HHAs referral concentration indices, and the average highest percent of patients admitted to a single PAC provider, were consistently higher for MJRLE patients than for CHF or sepsis patients. On average, Model 2 hospital EIs' MJRLE patients were admitted to a similar number of SNFs and HHAs, as was true for sepsis patients, while CHF patients were admitted to fewer SNFs but slightly more HHAs than MJRLE and sepsis patients. These absolute

---

[45] Total number of PAC providers in a market includes only those with at least one admission from a hospital of a patient in the particular clinical episode group during a six-month period. A Model 2 EI's patients can be admitted to a PAC facility in an adjacent CBSA. The number of PAC providers in a market varies by clinical episode type and time period.

[46] We did not estimate statistical significance of the difference in measure rates between the baseline and intervention period due to small sample size. Even at the national level (all EIs across all markets), the power analysis suggested the sample size was too small to reliably detect a statistically significant difference if one was present. We will re-evaluate this next year.

comparisons are not true within every market, however, as indicated in the individual markets of New York, Phoenix, and Providence.

Exhibit 44: Average Hospital-to-SNF Concentration Index for Major Joint Replacement of the Lower Extremity Episodes, Model 2, Q2 2011-Q3 2014

	Baseline				Intervention	
	Q4 2011 / Q1 2012	Q2 2012 / Q3 2012	Q4 2012 / Q1 2013	Q2 2013 / Q3 2013	Q4 2013 / Q1 2014*	Q2 2014 / Q3 2014
ALL MARKETS (101 EIs)						
Hospital-to-SNFs concentration index						
Mean	0.29	0.29	0.28	0.27	0.27	0.27
Median	0.20	0.19	0.19	0.21	0.19	0.21
25th Pctl	0.12	0.13	0.12	0.12	0.12	0.14
75th Pctl	0.38	0.41	0.37	0.32	0.38	0.33
Number of SNFs receiving patients, per Model 2 EI						
Mean	13.3	13.2	14.0	13.9	14.1	13.3
Highest percent of patients to one SNF						
Mean	41.1%	41.9%	40.3%	39.9%	39.3%	39.9%
NEW YORK MARKET (11 EIs)						
Mean concentration	0.17	0.18	0.16	0.20	0.16	0.18
SD	0.15	0.15	0.12	0.17	0.12	0.13
Number of SNFs receiving patients, per Model 2 EI						
Mean	26.5	25.2	26.7	25.7	28.3	26.5
Highest percent of patients to one SNF						
Mean	27.1%	30.3%	27.5%	29.3%	24.9%	27.5%
Number of SNFs receiving patients in the market						
Total	302	302	311	284	290	300
PHOENIX MARKET (7 EIs)						
Mean concentration	0.23	0.25	0.23	0.29	0.24	0.24
SD	0.16	0.12	0.12	0.17	0.12	0.08
Number of SNFs receiving patients, per Model 2 EI						
Mean	9	9	10	9	10.3	8.4
Highest percent of patients to one SNF						
Mean	34.9%	38.3%	33.7%	43.7%	37.9%	38.4%
Number of SNFs receiving patients in the market						
Total	73	74	68	69	76	66

	Baseline				Intervention	
	Q4 2011 / Q1 2012	**Q2 2012 / Q3 2012**	**Q4 2012 / Q1 2013**	**Q2 2013 / Q3 2013**	**Q4 2013 / Q1 2014***	**Q2 2014 / Q3 2014**
PROVIDENCE MARKET (4 EIs)						
Mean concentration	0.17	0.24	0.17	0.16	0.16	0.2
SD	0.12	0.21	0.19	0.12	0.10	0.2
Number of SNFs receiving patients, per Model 2 EI						
Mean	16.0	19.5	21.8	21.3	19.0	18.8
Highest percent of patients to one SNF						
Mean	32.2%	39.7%	27.4%	27.0%	32.4%	33.6%
Number of SNFs receiving patients in the market						
Total	83	92	88	95	86	84

Note: rates were calculated based on 101 Model 2 EIs with patients admitted to SNFs over all six periods. *A majority of Model 2 EIs joined BPCI in Q1 2014. *New York*: One Model 2 EI with patients admitted to SNFs joined BPCI in Q4 2013, 9 joined in Q1 2014, and 1 joined in Q2 2014. *Phoenix* and *Providence*: All Model 2 EIs in this table joined BPCI in Q1 2014.

Source: Lewin analysis of Medicare claims and enrollment data for episodes that began Q4 2011 through Q3 2014 for BPCI participants and a comparison group.

Exhibit 45: Average Hospital-to-HHAs Concentration Index for Major Joint Replacement of the Lower Extremity Episodes, Model 2, Q2 2011-Q3 2014

	Baseline				Intervention	
	Q4 2011 / Q1 2012	**Q2 2012 / Q3 2012**	**Q4 2012 / Q1 2013**	**Q2 2013 / Q3 2013**	**Q4 2013 / Q1 2014***	**Q2 2014 / Q3 2014**
ALL MARKETS (92 EIs)						
Hospital-to-HHAs concentration index						
Mean	0.45	0.41	0.47	0.43	0.44	0.46
Median	0.41	0.37	0.44	0.38	0.39	0.39
25th Pctl	0.28	0.26	0.29	0.26	0.25	0.28
75th Pctl	0.6	0.55	0.61	0.55	0.58	0.63
Number of HHAs receiving patients, per Model 2 EI						
Mean	7.6	8.2	7.8	8.1	8.0	7.5
Highest percent of patients to one HHA						
Mean	58.8%	54.9%	60.9%	57.1%	57.5%	58.9%
NEW YORK MARKET (6 EIs)						
Mean concentration	0.45	0.38	0.48	0.39	0.4	0.42
SD	0.22	0.12	0.17	0.20	0.19	0.19
Number of HHAs receiving patients, per Model 2 EI						
Mean	20.0	20.5	19.0	20.2	20.8	19.0
Highest percent of patients to one HHA						
Mean	60.2%	55.1%	66.9%	56.5%	57.8%	61.1%
Number of HHAs receiving patients in the market						
Total	98	100	94	90	86	86

	Baseline				Intervention	
	Q4 2011 / Q1 2012	**Q2 2012 / Q3 2012**	**Q4 2012 / Q1 2013**	**Q2 2013 / Q3 2013**	**Q4 2013 / Q1 2014***	**Q2 2014 / Q3 2014**
PHOENIX MARKET (7 EIs)						
Mean concentration	0.50	0.43	0.60	0.45	0.47	0.5
SD	0.19	0.14	0.23	0.14	0.29	0.27
Number of HHAs receiving patients, per Model 2 EI						
Mean	6.4	5	4.7	4.9	5.7	5
Highest percent of patients to one HHA						
Mean	63.8%	54.4%	72.1%	58.5%	57.7%	61.5%
Number of HHAs receiving patients in the market						
Total	44	34	39	38	39	42
PROVIDENCE MARKET (4 EIs)						
Mean concentration	0.37	0.34	0.45	0.28	0.34	0.41
SD	0.25	0.27	0.40	0.17	0.24	0.26
Number of HHAs receiving patients, per Model 2 EI						
Mean	6.5	8.0	7.0	8.5	7.8	7.0
Highest percent of patients to one HHA						
Mean	50.6%	44.0%	55.2%	41.3%	42.1%	53.4%
Number of HHAs receiving patients in the market						
Total	25	30	27	31	31	28

Note: rates were calculated based on 92 Model 2 EIs with patients admitted to HHAs over all six periods. *A majority of Model 2 EIs joined BPCI in Q1 2014. *New York*: One Model 2 EI with patients admitted to HHAs joined BPCI in Q4 2013, 4 joined in Q1 2014, and 1 joined in Q2 2014. *Phoenix* and *Providence*: All Model 2 EIs in this table joined BPCI in Q1 2014.

Source: Lewin analysis of Medicare claims and enrollment data for episodes that began Q4 2011 through Q3 2014 for BPCI participants and a comparison group.

2. Market share of Model 2 EIs

Exhibit 46 presents the trends for the market share of MJRLE episodes for Model 2 EIs across all markets. The average number of MJRLE episodes for Model 2 EIs in each period is also shown. The average market share of Model 2 hospital EIs participating in MJRLE episodes was 16 percent and did not increase after BPCI implementation. The same was true in the 3 markets we examined closely: the average MJRLE market share of BPCI EIs remained steady across the pre-BPCI and one year intervention periods, at about 5 percent for EIs in New York, 4 percent for EIs in Phoenix, and 8 percent for EIs in Providence. We see no evidence that BPCI EIs were capturing greater market share over time.

The time trends for the market shares of CHF and sepsis episodes are presented in **Appendix P**. Across all markets nationwide, the average CHF market share of Model 2 EIs did not increase between the pre-BPCI and one year intervention periods. The average sepsis market share for EIs increased from 15.4-15.7 percent in the pre-BPCI period to 16.3 percent in the one year intervention period, which coincided with a slight increase in the average number of sepsis episodes per EI. There was no change in average market shares of Model 2 hospital EIs for CHF or sepsis episodes in any of the three individual markets.

Exhibit 46: Average Market Share of Episode Initiators for Major Joint Replacement of the Lower Extremity Episodes, Model 2, Q2 2011-Q3 2014

	Baseline				Intervention	
	Q4 2011 / Q1 2012	Q2 2012 / Q3 2012	Q4 2012 / Q1 2013	Q2 2013 / Q3 2013	Q4 2013 / Q1 2014*	Q2 2014 / Q3 2014
ALL MARKETS (102 EIs)						
EI Market Share						
Mean	16.0%	16.2%	16.7%	16.4%	16.2%	16.4%
SD	19.4%	19.4%	19.9%	19.6%	19.0%	19.6%
Median	6.9%	8.3%	8.1%	7.7%	7.6%	8.0%
25th Pctl	3.6%	3.0%	2.9%	3.6%	3.7%	3.2%
75th Pctl	22.4%	22.5%	24.4%	22.4%	22.7%	22.4%
Number of MJRLE episodes per EI						
Mean	138.8	147.0	152.9	154.5	160.4	154.3
NEW YORK MARKET (11 EIs)						
EI Market Share						
Mean	4.7%	5.0%	4.8%	4.9%	5.0%	4.9%
SD	9.5%	10.1%	9.6%	9.8%	9.8%	9.9%
Number of MJRLE episodes per EI						
Mean	250.8	282.2	263.2	287.5	284.8	292.2
Number of hospitals with MJRLE admissions in the market						
Total	63	62	62	60	60	60
PHOENIX MARKET (7 EIs)						
EI Market Share						
Mean	3.7%	3.8%	3.8%	3.8%	3.8%	3.6%
SD	3.2%	3.2%	2.7%	2.5%	2.5%	2.3%
Number of MJRLE episodes per EI						
Mean	115.7	95.3	126.6	115	140.6	109.0
Number of hospitals with MJRLE admissions in the market						
Total	30	31	32	30	30	29
PROVIDENCE MARKET (4 EIs)						
EI Market Share						
Mean	6.9%	8.2%	8.7%	8.3%	8.5%	8.1%
SD	3.6%	5.8%	6.6%	7.3%	7.6%	7.0%
Number of MJRLE episodes per EI						
Mean	59.5	96.0	90.8	96.0	95.0	93.3
Number of hospitals with MJRLE admissions in the market						
Total	14	14	14	14	14	14

* A majority of Model 2 EIs joined BPCI in Q1 2014. *New York*: One Model 2 EI joined BPCI in Q4 2013, 9 joined in Q1 2014, and 1 EI admitting MJRLE episodes joined in Q2 2014. *Phoenix* and *Providence*: All Model 2 EIs joined BPCI in Q1 2014.

SD = standard deviation

Source: Lewin analysis of Medicare market share data from Q4 2011-Q3 2014.

D. Impact of BPCI on Orthopedic Surgery Episodes

1. Participants

During the first four quarters of the initiative, 82 hospitals, or 75% of hospital EIs in Model 2 participated in at least one orthopedic surgery episode.[47] There were 18,936 orthopedic surgery episodes initiated in Model 2-participating hospitals, accounting for 44% of all Model 2 episodes. Approximately 90% of these episodes were for major joint replacement of the lower extremity MS-DRGs.

Most of the hospitals participating in BPCI with these clinical episodes were urban facilities (95%) and 82% were non-profits. They were larger than the typical hospital (with an average of 334 beds, compared with 188 beds in non-participating hospitals) and likelier to be teaching hospitals. As to general characteristics, hospitals participating in these episodes were similar to other hospitals that were participating in BPCI.

2. Patient population characteristics

Medicare beneficiaries with an orthopedic surgery episode who were treated in a BPCI-participating hospital differed from all Medicare beneficiaries with the same type of episode with respect to age and Medicaid eligibility. The patients of BPCI providers tended to be older than all beneficiaries with the same MS-DRG, with a larger proportion aged 65-79 (64.4% vs. 58.9%; Exhibit 47). Among BPCI patients, 10.7% were eligible for Medicaid, compared with 15.1% of all Medicare beneficiaries with an orthopedic surgery episode. Both groups exhibited similar gender distributions and disability eligibility status.

Exhibit 47: Characteristics of BPCI Beneficiaries and All Medicare Beneficiaries with a Hospitalization for an Orthopedic Surgery MS-DRG, Model 2, Q4 2013-Q3 2014

Characteristics	Model 2 BPCI Beneficiaries with Orthopedic Surgery Episodes, Q4 2013 – Q3 2014 (N=17,934)		All Medicare Beneficiaries With Same MS-DRG Admission, Q3 2014	
	N	%	N	%
Age				
20-64	1,472	8.2%	22,694	11.5%
65-79	11,556	64.4%	115,834	58.9%
80+	4,906	27.4%	58,166	29.6%
Gender				
Female	11,843	66.0%	127,405	64.8%
Male	6,091	34.0%	69,289	35.2%

---

[47] The clinical episodes that compose the orthopedic surgery group are: amputation; double joint replacement of the lower extremity; hip & femur procedures except major joint; lower extremity and humerus procedure except hip, foot, femur; major joint replacement of the lower extremity; major joint replacement of the upper extremity; other knee procedures; removal of orthopedic devices; and revision of the hip or knee.

Characteristics	Model 2 BPCI Beneficiaries with Orthopedic Surgery Episodes, Q4 2013 – Q3 2014 (N=17,934)		All Medicare Beneficiaries With Same MS-DRG Admission, Q3 2014	
	N	%	N	%
Medicaid and Disability				
Percent Eligible for Medicaid	1,926	10.7%	29,707	15.1%
Percent Disability, no ESRD	1,946	10.9%	26,129	13.3%

Source: Lewin analysis of Medicare claims and enrollment data for episodes that began Q4 2011 through Q3 2014 for BPCI providers and for all admissions for the same MS-DRG to non-participating hospitals in Q3 2014.

3. *Change in Medicare standardized allowed payments*

Total standardized allowed payments during the anchor hospitalization and the 90-day post-discharge period declined more for episodes initiated by Model 2 hospital EIs than for those initiated in comparison hospitals. Although total standardized allowed payments for BPCI episodes were higher during the baseline period than they were for comparison episodes, payments fell more for BPCI episodes so that during the intervention period, BPCI episode payments were less than payments for the comparison episodes. As shown in Exhibit 48, average Medicare payments for the hospitalization and the care provided during the 90 days following discharge were estimated have declined $864 (or 3%) more for orthopedic surgery episodes initiated at BPCI providers than for orthopedic surgery episodes initiated by comparison hospitals. There was no indication that this reduction was due to providing some services prior to the anchor hospitalization or after the 90 day post-discharge period or by shifting to services that were excluded from the bundle. The change in Medicare allowed amounts before the bundle and after the PDP did not differ between episodes in BPCI providers and episodes in comparison providers. Similarly, there was no difference in the change in total standardized allowed payments for services outside of the bundled periods.

Exhibit 48: Impact of BPCI on Total Standardized Allowed Payment Inpatient Anchor Hospitalization and 90-day PDP for Orthopedic Surgery Episodes, Model 2, Q4 2011-Q3 2014

Measure	BPCI (N=17,672)		Comparison (N=17,688)		Diff-in-Diff estimate		
	Baseline	Intervention	Baseline	Intervention	Diff-in-Diff	LCI	UCI
Total Standardized Allowed Payment, Inpatient Anchor Hospitalization and 90-day PDP	$30,239	$28,232	$29,814	$28,670	-$864*	-$1,353	-$375

* Denotes statistical significance at the 5% level.

LCI = lower confidence interval at the 5% level; UCI = upper confidence interval at the 5% level.

Note: Baseline is defined as episodes that began Q4 2011 through Q3 2012. Intervention is defined as episodes that began Q4 2013 through Q3 2014. Sample sizes reflect number of episodes initiated during the intervention period that met inclusion criteria for the outcome.

Source: Lewin analysis of Medicare claims and enrollment data for episodes that began Q4 2011 through Q3 2014 for BPCI and comparison providers.

The greater decline in total allowed payments for BPCI episodes was due to changes in the use of PAC. Medicare IRF payments declined more for BPCI episodes than for comparison episodes, with a statistically significant difference of $459. Further, the decline in payments for the anchor

hospitalization and SNF care were also greater for BPCI episodes, although these were not statistically significant. The average SNF length of stay, however, was 1.3 days shorter during the intervention period than during the baseline for patients discharged from BPCI-participating hospitals, a statistically significant decline relative to the length of stay for the comparison group, which remained virtually unchanged.

The share of patients discharged to various PAC settings changed more for patients of BPCI-participating hospitals than for comparison hospitals. Prior to the BPCI intervention period, 64% of beneficiaries who received any PAC were discharged to institutional PAC (SNF, IRF, or long-term care hospitals). Under BPCI, this figure dropped to 57%. There is no evidence that reduced institutional PAC use for BPCI episodes was achieved by BPCI -participating hospitals changing to a less intensive mix of patients, as indicted by various measures of patient complexity (average hierarchical condition categories (HCC) score prior to the hospitalization, prior PAC use, or demographic characteristics).

These changes in use of PAC and discharge status are consistent with what we learned on site visits to BPCI-participating hospitals. According to interviews with hospital representatives, BPCI providers engaged in several types of activities to change patient and surgeon expectations about the use of PAC following a hip or knee replacement. Several hospital representatives reported that they strongly encouraged patients who were going to have a hip or knee replacement to attend a pre-operative educational program. Through these and other outreach activities, patients were advised of the advantages of a discharge home with home health care or outpatient physical therapy. Hospitals also reported working with orthopedic surgeons to engage them in supporting discharge home for their patients. Further, some orthopedic surgeons reported that shorter inpatient stays and more active rehabilitation were supported by changes in surgical techniques and improved pain management protocols.

4. Change in beneficiary quality of care

a. Claim and assessment-based measures

Early indications, based on claims and assessment data, are that there was no difference in the quality of care for Medicare beneficiaries receiving orthopedic surgery at BPCI-participating hospitals and those at comparison hospitals, with one exception with improvement in functioning. From the baseline to the intervention period, the change in hospital readmission rates, emergency department visits, and mortality following the anchor hospitalization for orthopedic surgery episodes was not statistically different in BPCI-participating hospitals and comparison providers (see Exhibit 49). Although the first Annual Report indicated that for episodes that began in the first quarter of the initiative the increase in emergency department use within 30 days for BPCI-participating hospitals was statistically significant relative to comparison hospitals, there are no longer any significant differences in emergency department use.

Exhibit 49: Impact of BPCI on Claims-based Quality Measures for Orthopedic Surgery Episodes, Model 2, Q4 2011-Q3 2014

Measure	BPCI (N=17,927)		Comparison (N=17,929)		Diff-in-Diff estimate		
	Baseline	Intervention	Baseline	Intervention	Diff-in-Diff	LCI	UCI
Unplanned Readmission Rate, within 30 days	6.4%	6.2%	6.9%	6.6%	0.0	-0.7	0.8
Unplanned Readmission Rate, within 90 days	10.9%	10.5%	11.3%	10.9%	-0.1	-1.1	0.9
Emergency Department Visit Rate, within 30 days	7.3%	7.7%	7.5%	7.7%	0.2	-0.6	1.0
Emergency Department Visit Rate, within 90 days	13.8%	14.4%	13.7%	14.2%	0.1	-1.0	1.2
Mortality, within 30 days	0.9%	1.0%	0.8%	1.0%	0.0	-0.3	0.2

* Denotes statistical significance at the 5% level.

LCI = lower confidence interval at the 5% level; UCI = upper confidence interval at the 5% level.

Note: Baseline is defined as episodes that began Q4 2011 through Q3 2012. Intervention is defined as episodes that began Q4 2013 through Q3 2014. Sample sizes reflect number of episodes initiated during the intervention period that met inclusion criteria for the outcome.

Source: Lewin analysis of Medicare claims and enrollment data for episodes that began Q4 2011 through Q3 2014 for BPCI and comparison providers.

Based on the patient assessment measures from PAC instruments, improvement in functioning was similar between beneficiaries whose episodes started in a BPCI-participating hospital and those whose episodes started in comparison hospitals. There was only one statistically significant finding: among beneficiaries who received their first PAC treatment at a HHA, the share of patients that exhibited improvement in upper body dressing declined from 94.3 percent to 93.8 percent among BPCI beneficiaries with orthopedic surgery episodes and increased from 93.9 percent to 95.0 percent among comparison beneficiaries who received their first PAC treatment at a HHA (Exhibit 50). This was a statistically significant and negative treatment effect of 1.6 percentage points.

Exhibit 50: Impact of BPCI on Assessment-based Quality Measures for Orthopedic Surgery Episodes, Model 2, Q4 2011 – Q2 2014

Measure	BPCI (N=3599)		Comparison (N=3030)		Diff-in-Diff estimate		
	Baseline	Intervention	Baseline	Intervention	Diff-in-Diff	LCI	UCI
Percent of HHA patients with improved upper body dressing	94.3%	93.8%	93.9%	95.0%	-1.6 pp*	-3.1 pp	-0.2 pp

* Denotes statistical significance at the 5% level.

LCI = lower confidence interval at the 5% level; UCI = upper confidence interval at the 5% level.

Note: Baseline is defined as episodes that began Q4 2011 through Q3 2012. Intervention is defined as episodes that began Q4 2013 through Q3 2014. Sample sizes reflect number of episodes initiated during the intervention period that met inclusion criteria for the outcome.

Source: Lewin analysis of patient assessment data for episodes that began Q4 2013 through Q2 2014 for BPCI and comparison providers.

b. Beneficiary survey measures

Exhibits 51 and 52 below report self-reported functional status improvement, functional status decline and patient care experience for beneficiaries in a MJRLE episode, which was the most

common orthopedic surgery episode. The proportion of respondents who improved was higher for BPCI respondents with MJRLE episodes than for the comparison respondents on two functional measures: "walk by yourself without resting" and "walk up and down 12 stairs" and statistically significant. Approximately 66% of the BPCI Model 2 respondents with MJRLE improved their ability to walk without resting as compared with 57.5% of comparison respondents. Similarly, 65.4% of the BPCI Model 2 respondents with MJRLE improved their ability to use stairs as compared with 57.9% among comparison respondents.

Exhibit 51: Improvement and Decline in Functional Status for BPCI and Comparison Survey Respondents for Major Joint Replacement of the Lower Extremity Episodes, Model 2, October and November 2014

Functional Measure	BPCI Rate	Comparison Rate	Treatment Effect
Improvement in bathing, dressing, using the toilet, or eating	85.0% (353)	84.5% (373)	0.5 [-4.3, 5.2]
Improvement in walking without rest	65.7% (351)	57.5% (374)	8.2* [1.6, 14.8]
Improvement in use of mobility device (i.e., less frequent)	60.8% (354)	63.5% (373)	-2.7 [-9.2, 3.8]
Improvement in using stairs	65.4% (353)	57.9% (366)	7.5* [0.9, 14.1]
Improvement in planning regular tasks	76.0% (355)	80.0% (377)	-4.0 [-9.4, 1.4]
Improvement in physical/emotional problems limiting social activities (i.e., less frequent)	75.8% (348)	72.0% (372)	3.8 [-2.4, 10.0]
Improvement in pain limiting regular activities (i.e., less frequent)	81.9% (355)	77.9% (378)	4.0 [-1.6, 9.7]
Decline in bathing, dressing, using the toilet, or eating	8.2% (353)	8.4% (373)	-0.1 [-3.9, 3.7]
Decline in walking without rest	14.4% (351)	15.7% (374)	-1.3 [-6.2, 3.5]
Decline in use of mobility device (i.e., more frequent)	26.7% (354)	24.2% (373)	2.5 [-3.3, 8.3]
Decline in using stairs	15.4% (353)	16.7% (366)	-1.3 [-6.0, 3.4]
Decline in planning regular tasks	13.7% (355)	9.5% (377)	4.2 [-0.1, 8.54]
Decline in physical/emotional problems limiting social activities (i.e., more frequent)	10.09% (348)	12.40% (372)	-2.3 [-6.0, 3.4]
Decline in pain limiting regular activities (i.e., more frequent)	7.04% (355)	8.49% (378)	-1.5 [-5.4, 2.5]

* Denotes statistical significance at the 5% level.

Notes: Treatment effect reported in percentage points. 95% confidence interval is in brackets. Survey results were weighted to adjust for sampling and nonresponse.

MJRLE refers to "major joint replacement of the lower extremity," *cardio* refers to "non-surgical cardiovascular", and *respiratory* refers to "non-surgical respiratory" episodes.

BPCI and comparison samples sizes for each question are reported in parentheses. The sample sizes vary between questions due to varying rates of item-response.

Source: Lewin analysis of Medicare beneficiary survey for Model 2 MJRLE BPCI and comparison group survey respondents.

For health care experience questions, BPCI respondents with MJRLE were less likely than comparison respondents to report that they agree or strongly agree that "medical staff clearly explained what follow-up appointments or treatment would be needed when prepared to go home." There were no statistically significant differences in responses to any other health care experience questions between BPCI and comparison respondents with MJRLE episodes.

Exhibit 52: Health Care Experience for BPCI and Comparison Survey Respondents for Major Joint Replacement of the Lower Extremity Episodes, Model 2, October and November 2014

Health Care Experience Measures	BPCI Rate	Comparison Rate	Treatment Effect
Thinking about all the care you received in the hospital before and afterwards, how often did you, your family, or your caregiver get conflicting advice from medical staff about your treatment? 1[response = Never]	81.2% (357)	83.0% (379)	-1.9 [-7.5, 3.8]
Thinking about all of the care you received in the hospital and afterwards, how often were the services you got appropriate for the level of care you needed? 1[response = Always]	67.6% (358)	70.1% (378)	-2.4 [-9.2, 4.4]
Thinking about when you left the hospital, were you discharged at the right time? 1[response = Yes]	89.9% (358)	93.3% (377)	-3.4 [-7.4, 0.7]
Thinking about when you left the hospital listed in the cover letter, the medical staff took your preferences and those of your family or your caregiver into account in deciding what health care services you should have after you left the hospital. 1[response = Agree/Strongly Agree]	95.9% (334)	95.0% (354)	0.8 [-2.4, 4.0]
Before you prepared to go home (or to someone else's home, or to an assisted living facility), you and your family or caregiver had a good understanding of how to take care of yourself. 1[response = Agree/Strongly Agree]	96.5% (342)	96.6% (357)	-0.2 [-2.9, 2.6]
Before you prepared to go home (or to someone else's home, or to an assisted living facility), medical staff clearly explained how to take your medications. 1[response = Agree/Strongly Agree]	95.1% (330)	96.0% (350)	-1.0 [-4.1, 2.2]
Before you prepared to go home (or to someone else's home, or to an assisted living facility), medical staff clearly explained what follow-up appointments or treatments would be needed. 1[response = Agree/Strongly Agree]	96.9% (332)	99.7% (341)	-2.8* [-4.7, -0.8]
Overall, since you returned home (or to someone else's home, or to an assisted living facility), you and your caregivers have been able to manage your health needs. 1[response = Agree/Strongly Agree]	99.7% (331)	98.6% (352)	1.0 [-0.3, 2.4]
Overall, how satisfied are you with your recovery since you left the hospital? 1[response = Extremely/Quite a Bit]	82.8% (341)	82.9% (363)	-0.1 [-5.8, 5.6]

* Denotes statistical significance at the 5% level.

Notes: Results are unadjusted. Treatment effect reported in percentage points. 95% confidence interval is in brackets. Survey results were weighted to adjust for sampling and nonresponse.

MJRLE refers to "major joint replacement of the lower extremity," *cardio* refers to "non-surgical cardiovascular", and *respiratory* refers to "non-surgical respiratory" episodes.

BPCI and comparison samples sizes for each question are reported in parentheses. The sample sizes vary between questions due to varying rates of item-response.

Source: Lewin analysis of Medicare beneficiary survey for Model 2 MJRLE BPCI and comparison group survey respondents.

c.   Technical expert panel

To help ensure that we were focusing on appropriate quality measures for beneficiaries with a hip or knee replacement and to understand where there may be potential vulnerabilities in quality, particularly given the shift away from institutional PAC, we convened a TEP to review preliminary data. The TEP was comprised of orthopedic surgeons, physical therapists, and other clinicians involved in orthopedic care. The panelists identified patients with comorbidities as more vulnerable to changes in quality of care and the need to monitor access to this surgery for more difficult patients, such as those with obesity or diabetes. They also agreed that home is the best site for recovery, although they stressed the need to ensure the safety of the home environment. They also indicated physical therapy is important in encouraging patients to move and in restoring normal movement patterns. In future analyses we will develop more information on any changes in physical therapy among BPCI patients relative to comparison patients, in conjunction with any changes in functional status. We will also be able to conduct more in depth analyses of quality of care for additional subpopulations, such as those with obesity or diabetes, who may be more vulnerable to changes in care and their access to care.

5.   Factors contributing to differences across BPCI hospital EIs

The change in total standardized allowed payment for Part A and B services during the anchor hospitalization and the 90-day PDP from baseline to intervention period among BPCI providers ranged from a net decline of $7,867 to a net increase of $4,163, with an average decline of $2,137 across all hospital EIs (Exhibit 53). Overall, 89% of providers experienced a decline in payments with 50% of the providers having a decline over $2,500.

Exhibit 53: Distribution of the Average Change in Total Standardized Allowed Payment for Part A and B Services during the Anchor Hospitalization and the 90-day PDP for Orthopedic Surgery Episodes, Model 2, Intervention relative to Baseline
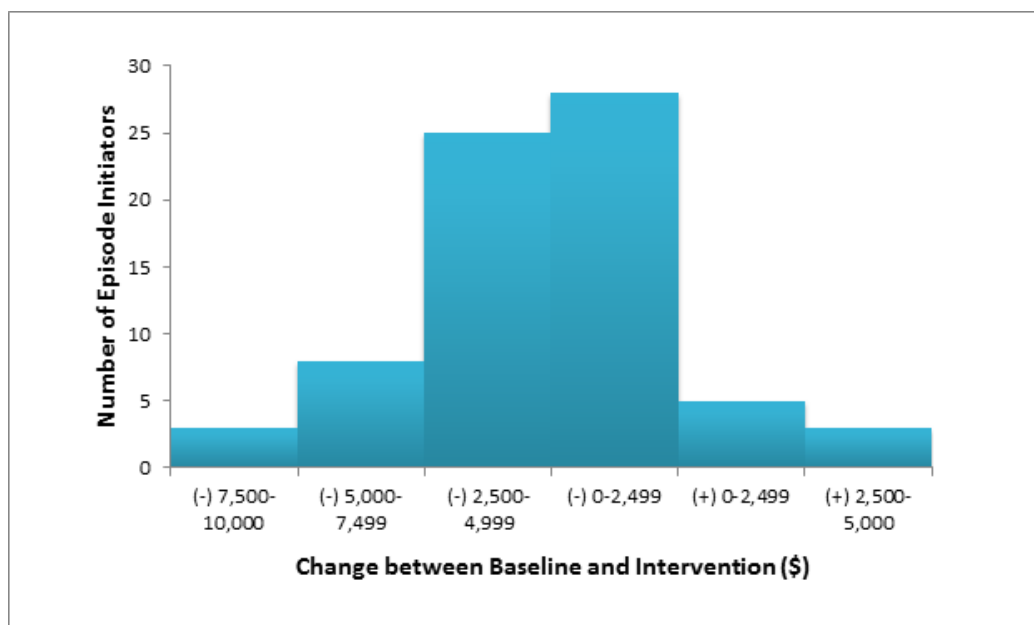
Note: Baseline is defined as episodes that began Q4 2011 through Q3 2012. Intervention is defined as episodes that began Q4 2013 through Q2 2014.

Source: Lewin analysis of Medicare claims and enrollment data for episodes that began Q4 2011 through Q2 2014 for BPCI providers.

To account for this variation, we estimated a regression model to determine the factors associated with the change in total standardized allowed payment from baseline to the intervention period after controlling for patient characteristics among BPCI episode initiators. In Exhibit 54 we present the predicted relative savings, calculated as the difference between the average savings from the baseline to intervention realized for providers with the characteristic and savings for providers without the characteristic, by provider-level and program characteristics. We found the overall reduction in total standardized allowed payment for Part A and B services during the anchor hospitalization and the 90-day PDP to be statistically significant ($2,211). We also found that BPCI hospitals that signed up to participate with the SNF 3-Day Waiver had an additional decline in payments ($899) relative to those who did not sign up to participate with this waiver. Since few BPCI providers that had signed up for the waiver actually used the waiver (see Section III.A.3.e above), we hypothesize that participating in the SNF 3-Day Waiver serves as a proxy for an effect undetermined by this analysis. Finally, we found that BPCI hospitals that were in markets with a higher concentration of SNF beds per 1,000 residents achieved a lower reduction in payments compared to BPCI providers in markets with a lower concentration of SNF beds ($1,672 vs. $2,573), that was statistically significant at 0.05. This finding is consistent with the information gathered through the Awardee interviews and site visits. BPCI participants stated that there is a challenge for providers to change the practice of discharging to a SNF as opposed to home with home health in markets that were saturated with SNFs or in instances where providers had direct relationships with SNFs.

Exhibit 54: Total Standardized Allowed Payment for Part A and B Services during the Anchor Hospitalization and the 90-day PDP by Provider and Program Characteristics for Orthopedic Surgery Episodes, Model 2, Intervention relative to Baseline

BPCI Characteristic	Risk-Adjusted Savings Intervention Period		Relative Difference in Savings		
	EI with characteristic present	EI without characteristic present	Difference	(95% Conf. Interval)	
All Episode initiators	$2,211.15*			$2,180.61	$2,241.98
By Case Mix Indicators					
Share of patients between 65 and 74 above median	$2,348.04	$2,024.63	$323.41	-$482.48	$1,129.30
Share of patients with disability above median	$2,155.88	$2,236.86	-$80.98	-$890.39	$728.44
By Program Characteristics					
Participating in Gainsharing	$2,170.55	$2,606.18	-$435.63	-$1,485.77	$614.52
Participating in SNF 3-Day waiver	$2,662.15	$1,764.31	$897.84*	$267.11	$1,528.56
Participating in Beneficiary incentives waiver	$2,243.76	$2,183.52	$60.24	-$606.08	$726.55
By Baseline Outcome					
Baseline Part A Payments, >75th %tile	$2,858.98	$2,060.33	$798.65	-$62.31	$1,659.61

	Risk-Adjusted Savings Intervention Period		Relative Difference in Savings		
BPCI Characteristic	**EI with characteristic present**	**EI without characteristic present**	**Difference**	**(95% Conf. Interval)**	
By Market Characteristics					
SNF Beds per 1000 above median	$1,672.30	$2,573.13	-$900.83*	-$1,640.66	-$160.98
Herfindahl Index above median	$1,866.50	$2,532.76	-$666.26	-$1,552.29	$219.75
Population in the market above median	$2,028.85	$2,379.90	-$351.05	-$1,291.83	$589.73

* Denotes statistical significance at the 5% level.

Note: The models above use 31,725 episodes from 72 BPCI hospitals. EIs with fewer than 25 episodes in the intervention period were excluded.

Note: Baseline is defined as episodes that began Q4 2011 through Q3 2012. Intervention is defined as episodes that began Q4 2013 through Q2 2014.

Source: Lewin analysis of Medicare claims and enrollment data for episodes that began Q4 2011 through Q2 2014 for BPCI providers.

a. *Total standardized allowed payment for Part A institutional services during the 90-day PDP*

The section above summarized the relative savings for total standardized allowed payment for Part A and B services during the anchor hospitalization and the 90-day PDP. In this section, we present the relative changes in total standardized allowed payment for Part A institutional services following the inpatient discharge from baseline to intervention period. The change from baseline to intervention period in total standardized allowed payment for Part A institutional services during the 90-day PDP among BPCI providers for orthopedic surgery episodes ranged from a net decline of $6,076 to a net increase of $3,290 with an average decline of $1,844 across all hospital EIs. Overall, 85% of EIs experienced a decline in payments with 33% of the providers having a decline over $2,500. The range of difference in payments between the baseline and intervention period is illustrated in Exhibit 55.

Exhibit 55: Distribution of the Average Change in Total Standardized Allowed Payment for Part A Institutional Services during the 90-day PDP for Orthopedic Surgery Episodes, Model 2, Intervention relative to Baseline
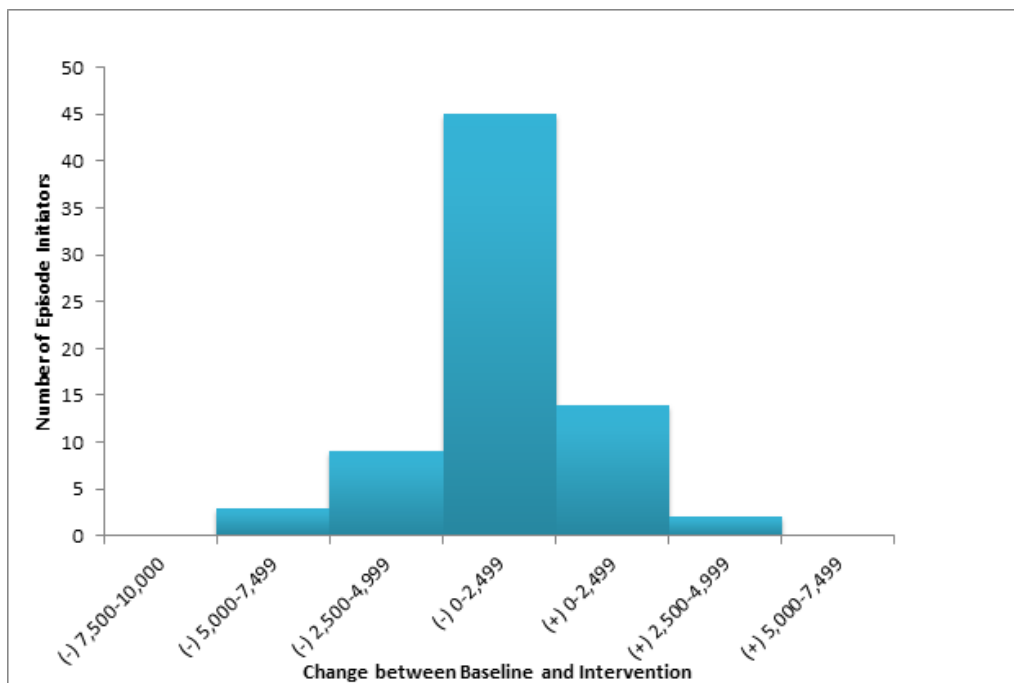
Note: Baseline is defined as episodes that began Q4 2011 through Q3 2012. Intervention is defined as episodes that began Q4 2013 through Q2 2014.

Source: Lewin analysis of Medicare claims and enrollment data for episodes that began Q4 2011 through Q2 2014 for BPCI providers.

In Exhibit 56 we present the predicted relative savings during the one year intervention period by the provider-level and program characteristics variables we included in the model. On average, BPCI providers reduced total standardized allowed payment for Part A institutional services during the 90-day PDP by $1,149. The characteristics associated with a greater decline in total standardized allowed payment for Part A institutional services during the 90-day PDP are consistent with the characteristics associated with a greater decline in total Medicare Part A and Part B payments during the anchor stay and 90 day PDP. Providers who signed up to participate with the SNF 3-Day Waiver and that had baseline payments greater than the 75th percentile had a greater decline in payments ($685 and $1,124, respectively) that was statistically significant. Providers that were in markets with a higher concentration of SNF beds achieved a lower reduction in payments compared to BPCI providers in markets with a lower concentration of SNF beds ($732.35 vs. $1430.59) that was statistically significant.

Exhibit 56: Average Change in Total Standardized Allowed Payment for Part A Institutional Services during the 90-day PDP by Provider and Program Characteristics for Orthopedic Surgery Episodes, Model 2, Intervention relative to Baseline

BPCI Characteristic	Risk-Adjusted Savings Intervention Period		Relative Difference in Savings		
	with characteristic present	without characteristic present	Difference	95% Conf. Interval	
All Episode initiators	$1,148.73*			$1,128.41	$1,169.05
By Case Mix Indicators					
Share of patients between 65 and 74 above median	$1,155.19	$1,139.99	$15.20	-$567.82	$598.22
Share of patients with disability above median	$1,127.00	$1,158.93	-$31.93	-$618.59	$554.73
By Program Characteristics					
Participating in Gainsharing	$1,082.06	$1,799.50	-$717.43	-$1,476.59	$41.72
Participating in SNF 3-Day waiver	$1,495.17	$810.19	$684.98*	$228.93	$1,141.03
Participating in Beneficiary incentives waiver	$971.06	$1,300.30	-$329.24	-$810.96	$152.49
By Baseline Outcome					
Baseline Part A Payments, >75th percentile	$2,063.49	$939.23	$1,124.26*	$503.05	$1,745.46
By Market Characteristics					
SNF Beds per 1000 above median	$732.35	$1,430.59	-$698.24*	-$1,233.04	-$163.44
Herfindahl Index above median	$1,054.54	$1,236.10	-$181.56	-$821.98	$458.85
Population in the market above median	$1,128.11	$1,167.94	-$39.83	-$718.46	$638.81

* Denotes statistical significance at the 5% level.

Note: The models above use 31,725 episodes from 72 BPCI hospitals. EIs with fewer than 25 episodes in the intervention period were excluded.

Note: Baseline is defined as episodes that began Q4 2011 through Q3 2012. Intervention is defined as episodes that began Q4 2013 through Q2 2014.

Source: Lewin analysis of Medicare claims and enrollment data for episodes that began Q4 2011 through Q2 2014 for BPCI providers.

b. 30-day unplanned readmission rate

There was no statistically significant decline in unplanned readmissions from the pre-BPCI period to the one year intervention period, and we were unable to detect statistically significant associations in the context of a regression model.

6. *Physician group practice episodes*

There were only three PGP EIs in the first year of the BPCI initiative, so we did not create a unique PGP comparison group.[48] Instead, we compared risk-adjusted PGP results with the observed (unadjusted) mean of all BPCI episodes in the same model and clinical episode group. This section compares the payment, utilization, and quality of care outcomes between the Model 2 PGP-initiated orthopedic surgery episodes and all Model 2 orthopedic surgery episodes. Given the imprecision in matching individual physicians to their PGP, these results should be viewed as preliminary and with caution.

The three PGP EIs in Model 2 initiated 889 orthopedic surgery episodes during the first year of BPCI (approximately 4.5% of all Model 2 orthopedic surgery episodes). The characteristics of the PGP-initiated orthopedic surgery clinical episode patients were similar to all Model 2 orthopedic surgery clinical episode patients with regards to age, prior health care utilization, and MS-DRG index. The PGP-initiated orthopedic surgery clinical episode patients were less likely to be eligible for Medicaid than all Model 2 orthopedic surgery clinical episode patients, both during the four quarters prior to BPCI (7.2% vs. 11.1%) and the first four quarters of the initiative (5.4% vs. 10.5%).

There was no statistical significance difference in the average total standardized allowed payment for Part A and B services during the anchor hospitalization and the 90-day PDP for PGP patients from the observed mean for all Model 2 orthopedic surgery episodes (see Exhibit 57). This is despite having higher total allowed payment for HHA services ($2,523) relative to all Model 2 orthopedic surgery episodes ($1,996) that was statistically significant at 0.05.

Exhibit 57: Total Allowed Payment Amount Inpatient Anchor Hospitalization and 90-day PDP for PGP-Initiated Orthopedic Surgery Episodes and All Orthopedic Surgery Episodes, Model 2, Q4 2013-Q3 2014

Measure	PGP (N=836)			All Model 2 episodes (N=18,598)	Model 2 PGP Risk-adjusted Mean, vs. All Model 2, Observed Mean
	Risk-adjusted mean	LCI	UCI	Observed Mean	
Total Allowed Payment Amount, Inpatient Anchor Hospitalization and 90-day PDP	$28,029	$26,990	$29,068	$27,945	$84

* Denotes statistical significance at the 5% level.

LCI = lower confidence interval at the 5% level; UCI = upper confidence interval at the 5% level.

Source: Lewin analysis of Medicare claims and enrollment data for episodes that began Q4 2013 through Q3 2014 for BPCI PGP and hospital providers.

The higher spending on HHA services among the PGP-initiated orthopedic surgery episodes was driven by the fact that these patients were more likely to be discharged to an HHA than all Model 2 orthopedic surgery patients. While 52.5% of patients in PGP-initiated orthopedic surgery episodes were discharged to a HHA, only 36.2% of all Model 2 orthopedic surgery episodes were

---

[48] Please note that the data linking individual physicians with their PGP contained errors, so these results should be viewed with caution.

similarly discharged. The patients in PGP-initiated orthopedic surgery episodes were therefore less likely to be discharged home with no home health (9% vs. 18% of all Model 2 orthopedic surgery episodes) and less likely to be discharged to a SNF (32.5% vs. 36.9% of all Model 2 orthopedic surgery episodes). This finding was statistically significant.

The quality of care for BPCI patients with PGP-initiated orthopedic surgery episodes was similar to that of all Model 2 orthopedic surgery patients as measured through various claim-based measures and PAC assessments. As shown in Exhibit 58, hospital readmission rates, emergency department visits, and mortality following the anchor hospitalization were not statistically different for PGP patients relative to all BPCI patients in Model 2 orthopedic surgery episodes. (Hospital readmission rates and emergency department visits in the 90-day PDP were not produced for this Model 2 PGP clinical group.)

Exhibit 58: Claims-based Quality Measures for PGP-Initiated Orthopedic Surgery Episodes and All Orthopedic Surgery Episodes, Model 2, Q4 2013-Q3 2014

Measure	PGP (N=846)			All Model 2 episodes (N=18,864)	Model 2 PGP Risk-adjusted Mean, vs. All Model 2, Observed Mean
	Risk-adjusted mean	LCI	UCI	Observed Mean	
Unplanned readmission rate, 30-day PDP	6.0%	4.3%	7.6%	6.1%	-0.1
ED use, 30-day PDP	7.9%	6.1%	9.7%	7.5%	0.4
All-cause 30-day mortality rate	0.8%	0.2%	1.3%	0.9%	-0.2

* Denotes statistical significance at the 5% level.

LCI = lower confidence interval at the 5% level; UCI = upper confidence interval at the 5% level.

Source: Lewin analysis of Medicare claims and enrollment data for episodes that began Q4 2013 through Q3 2014 for BPCI PGP and hospital providers.

E.   Impact of BPCI on Cardiovascular Surgery

1.   Participants

During the first four quarters of the initiative, 30 hospitals, or 27% of hospital EIs in Model 2, participated in at least one cardiovascular surgery episode.[49] There were 2,859 cardiovascular surgery episodes initiated in Model 2-participating hospitals, accounting for 7% of all Model 2 episodes. Approximately 32% of these episodes were for cardiac valve MS-DRGs and 28% were for coronary artery bypass graft MS-DRGs. Each of the hospitals participating in BPCI with these clinical episodes was an urban facility and 93% were non-profit organizations. They were larger than the typical hospital (with an average of 463 beds, compared with 188 beds in non-participating hospitals) and likelier to be teaching hospitals.

---

[49] The clinical episodes that compose the cardiovascular surgery group are: AICD generator or lead; cardiac defibrillator; cardiac valve; coronary artery bypass graft; major cardiovascular procedure; other vascular surgery; pacemaker; pacemaker device replacement or revision; and percutaneous coronary intervention.

2. Patient population characteristics

Beneficiaries who were treated in a hospital participating in Model 2 for a cardiovascular surgery episode were similar to all Medicare beneficiaries treated in the same MS-DRG. Both groups exhibit similar age, gender, dual eligibility status, and Medicare eligibility due to disability (see Exhibit 59).

Exhibit 59: Characteristics of BPCI Beneficiaries and All Medicare Beneficiaries with a Hospitalization for a Cardiovascular Surgery MS-DRG, Model 2, Q4 2013-Q3 2014

Characteristics	Model 2 BPCI Beneficiaries with Cardiovascular Surgery Episodes, Q4 2013 – Q3 2014 (N=2,718)		All Medicare Beneficiaries With Same MS-DRG Admission, Q3 2014 (N=62,576)	
	N	%	N	%
Age				
20-64	247	9.1%	8,023	12.8%
65-79	1,578	58.1%	34,213	54.7%
80+	893	32.9%	20,340	32.5%
Gender				
Female	1,083	39.8%	26,340	42.1%
Male	1,635	60.2%	36,236	57.9%
Medicaid and Disability				
Percent Eligible for Medicaid	445	16.4%	11,177	17.9%
Percent Disability, no ESRD	327	12.0%	6,931	11.1%

Source: Lewin analysis of Medicare claims and enrollment data for episodes that began Q4 2011 through Q3 2014 for BPCI providers and for all admissions for the same MS-DRG to non-participating hospitals in Q3 2014.

3. Change in Medicare standardized allowed payments

There were statistically significant changes in utilization of PAC for BPCI episodes relative to comparison group episodes. While these changes did not drive a statistically significant difference in total standardized Part A and B payments during the anchor hospitalization and the 90-day PDP, the standardized allowed payment included in the bundle did have a statistically significant decrease for the 30-day episodes. The proportion of beneficiaries who were discharged to institutional PAC among those who received any PAC decreased from 55% in baseline to 44% in the intervention period for BPCI beneficiaries, compared with a decrease from 47.2% to 46.2% for beneficiaries treated by comparison providers. This resulted in a 10 percentage point greater decrease among BPCI beneficiaries relative to the comparison group. At the same time, the number of HH visits increased for BPCI beneficiaries from 15.8 to 16.6 and decreased from 16.6 to 15.9 visits for the comparison group. This resulted in a statistically significant increase of 1.5 visits for BPCI relative to the comparison group. The average total standardized allowed payment for Part A and B services during the anchor hospitalization and the 90-day PDP for beneficiaries treated by BPCI providers declined from baseline to the intervention periods, but this reduction was not statistically significant relative to the change for beneficiaries treated by comparison providers (see Exhibit 60). However, the total standardized allowed payment included in the bundle for 30 day episodes with PAC use did have a statistically significant decrease for beneficiaries treated by BPCI providers relative to the comparison group (-$4,149). Approximately 13% of cardiovascular surgery episodes were 30-day episodes. The remaining 87% of the episodes were 90 day episodes; the total allowed payment included in the bundle for 90 day

episodes was not statistically significant. We will continue to monitor the within bundle allowed payment amount in future quarters of the initiative. There were no indications that BPCI participating providers shifted services outside of the bundle definition or bundle period.

Exhibit 60: Impact of BPCI on Total Standardized Allowed Payment Inpatient Anchor Hospitalization and 90-day PDP for Cardiovascular Surgery Episodes, Model 2, Q4 2011-Q3 2014

Measure	BPCI (N=2,641)		Comparison (N=2,653)		Diff-in-Diff estimate		
	Baseline	Intervention	Baseline	Intervention	Diff-in-Diff	LCI	UCI
Total Standardized Allowed Payment, Inpatient Anchor Hospitalization and 90-day PDP	$46,805	$46,282	$46,272	$46,628	-$880	-$2,584	$825

* Denotes statistical significance at the 5% level.

LCI = lower confidence interval at the 5% level; UCI = upper confidence interval at the 5% level.

Note: Baseline is defined as episodes that began Q4 2011 through Q3 2012. Intervention is defined as episodes that began Q4 2013 through Q3 2014. Sample sizes reflect number of episodes initiated during the intervention period that met inclusion criteria for the outcome.

Source: Lewin analysis of Medicare claims and enrollment data for episodes that began Q4 2011 through Q3 2014 for BPCI and comparison providers.

4. *Change in beneficiary quality of care*

The emergency department visit rate during the 30-day PDP increased significantly for BPCI beneficiaries relative to the comparison group (see Exhibit 61). The emergency department visit rate increased from 9.3% to 12.0% for BPCI beneficiaries during the 30-day PDP, compared with 10.8 to 11.3% for comparison beneficiaries. The change in emergency department visit rate during the 90-day PDP, however, was not statistically different between BPCI and the comparison group beneficiaries. BPCI patients with cardiovascular surgery episodes had a statistically significant increase in mortality rates from baseline to intervention relative to the patients treated by the matched comparison group. The mortality rate increased from 1.6% to 1.9% for BPCI patients. Mortality rates were higher during the baseline period for comparison providers, and declined more, from 2.1% to 1.4%. (Given the importance of this negative outcome, more recent results based on a more refined comparison group were examined while this report was in draft form. The updated analysis indicates that this negative outcome did not persist.)[50] The change in hospital readmission rates was not statistically different for patients of BPCI participants and comparison providers. There were no statistically significant changes in assessment-based quality measures from baseline to intervention period between BPCI and

---

[50] More recent results based on data from Q4 2013 to Q2 2015 and a refined comparison group indicate that, for three of the clinical episodes in the cardiovascular surgery clinical episode group, the 30-day mortality rate for BPCI beneficiaries did not have a statistically significant change (at the 5%, 10%, or 20% levels) relative to comparison beneficiaries. The 30-day mortality DiD estimate for coronary artery bypass graft (CABG) episodes was 0.05 (95% CI: -1.05, 1.16), for percutaneous coronary intervention (PCI) episodes was 0.65 (95% CI : -0.80, 2.10), and for cardiac valve episodes was -0.26 (95% CI: -2.46, 1.95). The 90-day mortality DiD estimates were also not statistically significant (at the 5%, 10%, or 20% levels). The 90-day mortality DiD for CABG episodes was 0.77 (95% CI: -0.81, 2.36), for PCI episodes was 1.02 (95% CI: -0.57, 2.61), and for cardiac valve episodes was 0.31 (95% CI: -1.32, 1.94).

comparison group beneficiaries with cardiovascular surgery episodes who received their first PAC treatment at a SNF, a HHA or an IRF.

Exhibit 61: Impact of BPCI on Claims-based Quality Measures for Cardiovascular Surgery Episodes, Model 2, Q4 2011-Q3 2014

Measure	BPCI (N=2,712)		Comparison (N=2,714)		Diff-in-Diff estimate		
	Baseline	Intervention	Baseline	Intervention	Diff-in-Diff	LCI	UCI
Unplanned Readmission Rate, within 30 days	15.7%	15.4%	14.9%	16.0%	-1.3	-3.7	1.2
Unplanned Readmission Rate, within 90 days	25.0%	24.2%	24.8%	24.7%	-0.7	-3.2	1.8
Emergency Department Visit Rate, within 30 days	9.3%	12.0%	10.8%	11.3%	2.2*	0.1	4.4
Emergency Department Visit Rate, within 90 days	17.8%	20.9%	19.7%	22.3%	0.4	-2.3	3.0
Mortality, within 30 days	1.6%	1.9%	2.1%	1.4%	0.9*	0.0	1.8

* Denotes statistical significance at the 5% level.

LCI = lower confidence interval at the 5% level; UCI = upper confidence interval at the 5% level.

Note: Baseline is defined as episodes that began Q4 2011 through Q3 2012. Intervention is defined as episodes that began Q4 2013 through Q3 2014. Sample sizes reflect number of episodes initiated during the intervention period that met inclusion criteria for the outcome.

Source: Lewin analysis of Medicare claims and enrollment data for episodes that began Q4 2011 through Q3 2014 for BPCI and comparison providers.

F.   Impact of BPCI on Spinal Surgery

1.   Participants

Twenty Model 2 hospital EIs, or 18% of hospital EIs in Model 2, participated in at least one spinal surgery clinical episode during the first four quarters of the initiative. There were 966 spinal surgery clinical episodes initiated in Model 2 participating hospitals, accounting for 2% of all Model 2 episodes. Approximately 59% of these episodes were for spinal fusion MS-DRGs. Each of the hospitals participating in BPCI with these clinical episodes was an urban facility and 85% were non-profit organizations. They were larger than the typical hospital (with an average of 367 beds, compared with 188 beds in non-participating hospitals) and likelier to be teaching hospitals.

2.   Patient population characteristics

Beneficiaries who were treated in a hospital participating in Model 2 for a spinal surgery episode were similar to all Medicare beneficiaries treated in the same MS-DRG. Both groups exhibit similar gender and eligibility for public programs (Exhibit 62). The major difference between the BPCI and non-BPCI groups was related to age. The patients of BPCI providers were older than all beneficiaries with the same MS-DRG, with a larger proportion aged 80 and older (14.9% vs. 10.5%).

Exhibit 62: Characteristics of BPCI Beneficiaries and All Medicare Beneficiaries with a Hospitalization for a Spinal Surgery MS-DRG, Model 2, Q4 2013 - Q3 2014

Characteristics	Model 2 Spinal Surgery Episodes BPCI Beneficiaries, Q4 2013-Q3 2014 (N=898)		All Medicare Beneficiaries with same MS-DRG admission, Q3 2014 (N=33,275)	
	N	%	N	%
Age				
20-64	203	22.6%	8,023	24.1%
65-79	561	62.5%	21,773	65.4%
80+	134	14.9%	3,479	10.5%
Gender				
Female	508	56.6%	18,919	56.9%
Male	390	43.4%	14,356	43.1%
Medicaid and Disability				
Percent Eligible for Medicaid	148	16.5%	5,732	17.2%
Percent Disability, no ESRD	242	26.9%	9,290	27.9%

Source: Lewin analysis of Medicare claims and enrollment data for episodes that began Q4 2011 through Q3 2014 for BPCI providers and for all admissions for the same MS-DRG to non-participating hospitals in Q3 2014.

3. *Change in Medicare standardized allowed payments*

Total standardized allowed payment for Part A and B services during the anchor hospitalization and the 90-day PDP increased more for BPCI beneficiaries than for beneficiaries in the comparison group. As shown in Exhibit 63, average payments for patients treated by BPCI providers in spinal surgery episodes increased from $42,202 in baseline to $46,234 in the intervention period. This was $3,477 more relative to the increase for patients of comparison providers ($40,691 to $41,246). Part B payments for imaging and lab had a statistically significant increase of $53 more for BPCI beneficiaries than comparison beneficiaries. Standardized Part A and B payments during the anchor stay had greater increases among BPCI beneficiaries relative to the comparison group, although these changes were not statistically significant. For example, standardized allowed Part A payment for the anchor hospitalization increased $2,285 more for BPCI beneficiaries than comparison beneficiaries. For BPCI beneficiaries the anchor hospitalization increased from $24,233 in the baseline to $27,270 in the intervention period compared to $23,735 in the baseline and $24,487 in the intervention period for comparison beneficiaries. Part B payments for procedures during the anchor stay increased by $242 more for BPCI beneficiaries than comparison beneficiaries. In addition, while not risk adjusted due to limited sample size, SNF and IRF payment increases were also greater for BPCI beneficiaries ($412 and $400 respectively). There were no statistically significant differences in PAC utilization that drove this increase in payments although the number of institutional days during the 90-day PDP increased 3.4 days more for BPCI beneficiaries relative to comparison beneficiaries. These differences, although not statistically significant, may suggest differences in treatment between BPCI and comparison providers for spinal surgery episodes.

Exhibit 63: Impact of BPCI on Total Standardized Allowed Payment Inpatient Anchor Hospitalization and 90-day PDP for Spinal Surgery Episodes, Model 2, Q4 2011-Q3 2014

Measure	BPCI (N=881)		Comparison (N=877)		Diff-in-Diff estimate		
	Baseline	Intervention	Baseline	Intervention	Diff-in-Diff	LCI	UCI
Total Standardized Allowed Payment, Inpatient Anchor Hospitalization and 90-day PDP	$42,202	$46,234	$40,691	$41,246	$3,477*	$228	$6,726

* Denotes statistical significance at the 5% level.

LCI = lower confidence interval at the 5% level; UCI = upper confidence interval at the 5% level.

Note: Baseline is defined as episodes that began Q4 2011 through Q3 2012. Intervention is defined as episodes that began Q4 2013 through Q3 2014. Sample sizes reflect number of episodes initiated during the intervention period that met inclusion criteria for the outcome.

Source: Lewin analysis of Medicare claims and enrollment data for episodes that began Q4 2011 through Q3 2014 for BPCI and comparison providers.

4. *Change in beneficiary quality of care*

The quality of care for BPCI beneficiaries with spinal surgery episodes was similar to that of beneficiaries treated by non-BPCI providers, as measured through various claim-based measures and PAC assessments, with one exception. As shown in Exhibit 64, mortality declined from 1.2% to 0.2% for BPCI beneficiaries with spinal surgery episodes and increased from 0.5% to 1.2% for beneficiaries with episodes in comparison providers. This resulted in a statistically significant difference of 1.6 percentage points. (Because of the importance of this outcome, more recent results based on a more refined comparison group matching approach were examined while this report was in draft form. The updated analysis indicates that this positive outcome did not appear to persist.)[51] The change in unplanned hospital readmission rates and emergency department visits was not statistically different for patients of BPCI providers and comparison providers. There were no statistically significant differences in assessment-based quality measures from baseline to intervention period between BPCI and comparison group beneficiaries with spinal surgery episodes who received their first PAC treatment at a SNF, a HHA or an IRF.

---

[51] More recent results, based on data from Q4 2013 to Q2 2015, indicate that BPCI beneficiaries with spinal fusion episodes (a clinical episode in the spinal surgery clinical episode group) did not have a statistically significant change (at the 5%, 10%, or 20% levels) in 90-day mortality rates relative to comparison beneficiaries. The 90-day mortality DiD estimate for spinal fusion episodes was 0.25 (95% CI: -1.10, 1.60). For Q2 2015, it was not possible to produce risk-adjusted 30-day mortality results because of the low occurrence of this event.

Exhibit 64: Impact of BPCI on Claims-based Quality Measures for Spinal Surgery Episodes, Model 2, Q4 2011-Q3 2014

Measure	BPCI (N=895)		Comparison (N=897)		Diff-in-Diff estimate		
	Baseline	Intervention	Baseline	Intervention	Diff-in-Diff	LCI	UCI
Unplanned Readmission Rate, within 30 days	8.9%	9.2%	9.8%	9.3%	0.7	-3.2	4.6
Unplanned Readmission Rate, within 90 days	13.9%	15.0%	14.2%	13.7%	1.6	-3.4	6.7
Emergency Department Visit Rate, within 30 days	10.9%	13.3%	9.1%	9.0%	2.5	-0.7	5.7
Emergency Department Visit Rate, within 90 days	18.9%	21.7%	17.2%	17.8%	2.2	-2.5	6.8
Mortality, within 30 days	1.2%	0.2%	0.5%	1.2%	-1.6*	-2.7	-0.4

* Denotes statistical significance at the 5% level.

LCI = lower confidence interval at the 5% level; UCI = upper confidence interval at the 5% level.

Note: Baseline is defined as episodes that began Q4 2011 through Q3 2012. Intervention is defined as episodes that began Q4 2013 through Q3 2014. Sample sizes reflect number of episodes initiated during the intervention period that met inclusion criteria for the outcome.

Source: Lewin analysis of Medicare claims and enrollment data for episodes that began Q4 2011 through Q3 2014 for BPCI and comparison providers.

G.   Impact of BPCI on Non-Surgical and Surgical GI

1.   Participants

During the first four quarters of the initiative, 17 hospitals, or 15% of hospital EIs in Model 2, participated in at least one non-surgical and surgical GI episode.[52] There were 1,464 non-surgical and surgical GI episodes initiated in Model 2-participating hospitals, accounting for 3% of all Model 2 episodes. Approximately 44% of these episodes were for gastrointestinal hemorrhage MS-DRGs. Each of the hospitals participating in BPCI with these clinical episodes was in an urban facility and 94% were non-profits. They were larger than the typical hospital (with an average of 332 beds, compared with 188 beds in non-participating hospitals) and were likelier to be teaching hospitals.

2.   Patient population characteristics

Beneficiaries who were treated in a hospital participating in Model 2 for a non-surgical and surgical GI episode were different from all Medicare beneficiaries treated in the same MS-DRG. The major differences between BPCI and non-BPCI groups were related to age and gender. The patients of BPCI providers were older than all beneficiaries with the same MS-DRG, with a larger proportion aged 80 and older (46.9% v. 38.6%; Exhibit 65). Among BPCI beneficiaries, 61.4% were female, compared with 55.8% of all Medicare beneficiaries. Both groups were similar with respect to dual eligibility status and Medicare eligibility due to disability.

---

[52] The clinical episodes that compose the non-surgical and surgical GI group are: esophagitis, gastroenteritis and other digestive disorders; gastrointestinal hemorrhage; gastrointestinal obstruction; and major bowel procedure.

Exhibit 65: Characteristics of BPCI Beneficiaries and All Medicare Beneficiaries with a Hospitalization for a Non-Surgical and Surgical GI MS-DRG, Model 2, Q4 2013-Q3 2014

Characteristics	Model 2 BPCI Beneficiaries with Non-Surgical and Surgical GI Episodes, Q4 2013 – Q3 2014 (N=1,400)		All Medicare Beneficiaries With Same MS-DRG Admission, Q3 2014 (N=100,465)	
	N	%	N	%
Age				
20-64	141	10.1%	13,715	13.7%
65-79	602	43.0%	47,931	47.7%
80+	657	46.9%	38,819	38.6%
Gender				
Female	860	61.4%	56,020	55.8%
Male	540	38.6%	44,445	44.2%
Medicaid and Disability				
Percent Eligible for Medicaid	250	17.9%	22,281	22.2%
Percent Disability, no ESRD	168	12.0%	14,512	14.4%

Source: Lewin analysis of Medicare claims and enrollment data for episodes that began Q4 2011 through Q3 2014 for BPCI providers and for all admissions for the same MS-DRG to non-participating hospitals in Q3 2014.

3.   Change in Medicare standardized allowed payment

There were no statistically significant differences in standardized allowed payments between BPCI and comparison group episodes. The average total standardized allowed payment for Part A and B services during the anchor hospitalization and the 90-day PDP for beneficiaries treated by BPCI providers declined from baseline to intervention periods, but this reduction was not statistically significant relative to the change for beneficiaries treated by matched comparison providers (see Exhibit 66). Furthermore, there were no statistically significant differences in the change for the allowed amount for any Part A or Part B service. There were also no statistically significant differences in the change in PAC use – both the first PAC setting and the intensity of PAC services (i.e. length of stay or number of home health visits) - relative to the episodes from matched comparison providers. There were no indications that these results were achieved by shifting services outside the bundle period or not covered in the bundle definition.

Exhibit 66: Impact of BPCI on Total Standardized Allowed Payment Inpatient Anchor Hospitalization and 90-day PDP for Non-Surgical and Surgical GI Episodes, Model 2, Q4 2011 – Q3 2014

Measure	BPCI (N=1,377)		Comparison (N=1,375)		Diff-in-Diff estimate		
	Baseline	Intervention	Baseline	Intervention	Diff-in-Diff	LCI	UCI
Total Standardized Allowed Payment, Inpatient Anchor Hospitalization and 90-day PDP	$25,723	$24,289	$25,045	$24,295	-$684	-$2,932	$1,564

* Denotes statistical significance at the 5% level.

LCI = lower confidence interval at the 5% level; UCI = upper confidence interval at the 5% level.

Note: Baseline is defined as episodes that began Q4 2011 through Q3 2012. Intervention is defined as episodes that began Q4 2013 through Q3 2014. Sample sizes reflect number of episodes initiated during the intervention period that met inclusion criteria for the outcome.

Source: Lewin analysis of Medicare claims and enrollment data for episodes that began Q4 2011 through Q3 2014 for BPCI and comparison providers.

4.   *Change in beneficiary quality of care*

The quality of care for BPCI beneficiaries with non-surgical and surgical GI episodes was similar to that of beneficiaries treated by non-BPCI providers, as measured through various claim-based measures. As seen in Exhibit 67, the change in hospital readmission rates, emergency department visits, and mortality following the anchor hospitalization for non-surgical and surgical GI episodes was not statistically different for patients of BPCI providers and comparison providers.

Exhibit 67: Impact of BPCI on Claims-based Quality Measures for Non-Surgical and Surgical GI Episodes, Model 2, Q4 2011-Q3 2014

Measure	BPCI (N=1,386)		Comparison (N=1,389)		Diff-in-Diff estimate		
	Baseline	Intervention	Baseline	Intervention	Diff-in-Diff	LCI	UCI
Unplanned Readmission Rate, within 30 days	17.2%	14.7%	16.0%	14.3%	-0.8	-4.3	2.7
Unplanned Readmission Rate, within 90 days	27.7%	26.9%	27.9%	25.6%	1.4	-2.7	5.5
Emergency Department Visit Rate, within 30 days	9.8%	11.1%	10.4%	10.3%	1.4	-1.7	4.5
Emergency Department Visit Rate, within 90 days	21.1%	20.4%	20.3%	19.5%	0.1	-3.2	3.5
Mortality, within 30 days	4.3%	3.8%	4.3%	3.4%	0.3	-1.8	2.4

* Denotes statistical significance at the 5% level.

LCI = lower confidence interval at the 5% level; UCI = upper confidence interval at the 5% level.

Note: Baseline is defined as episodes that began Q4 2011 through Q3 2012. Intervention is defined as episodes that began Q4 2013 through Q3 2014. Sample sizes reflect number of episodes initiated during the intervention period that met inclusion criteria for the outcome.

Source: Lewin analysis of Medicare claims and enrollment data for episodes that began Q4 2011 through Q3 2014 for BPCI and comparison providers.

Among patients with non-surgical and surgical GI episodes who received their first PAC treatment at a HHA, there was a statistically significant reduction in the share of BPCI beneficiaries who improved in upper body dressing (12.0 percentage points) and in lower body dressing (14.3 percentage points), relative to the comparison group (Exhibit 68). For both measures, the statistically significant negative effect of the BPCI intervention was caused by lower rates from baseline to intervention among the BPCI patients and higher rates among comparison group patients during the same period.

Exhibit 68: Impact of BPCI on Assessment-based Quality Measures for Non-surgical and Surgical GI Episodes, Model 2, Q4 2011 – Q2 2014

Measure	BPCI (N=102)		Comparison (N=117)		Diff-in-Diff estimate		
	Baseline	Intervention	Baseline	Intervention	Diff-in-Diff	LCI	UCI
HHA patients with improved upper body dressing	82.4%	75.6%	71.2%	76.5%	-12.0pp*	-24.1 pp	-0.03 pp
HHA patients with improved lower body dressing	83.2%	72.9%	71.8%	75.9%	-14.3 pp*	-23.6 pp	-5.2 pp

* Denotes statistical significance at the 5% level.

LCI = lower confidence interval at the 5% level; UCI = upper confidence interval at the 5% level.

Note: Baseline is defined as episodes that began Q4 2011 through Q3 2012. Intervention is defined as episodes that began Q4 2013 through Q2 2014. Sample sizes reflect number of episodes initiated during the intervention period that met inclusion criteria for the outcome.

Source: Lewin analysis of patient assessment data for episodes that began Q4 2013 through Q2 2014 for BPCI and comparison providers.

H. Impact of BPCI on Non-Surgical Cardiovascular

1. Participants

During the first four quarters of the initiative, 41 hospitals, or 37% of hospital EIs in Model 2, participated in at least one non-surgical cardiovascular episode.[53] There were 6,661 non-surgical cardiovascular episodes initiated in Model 2-participating hospitals, accounting for 16% of all Model 2 episodes. Approximately 72% of these episodes were for congestive heart failure. Hospitals participating in BPCI with these clinical episodes were predominantly urban (95%), non-profits (95%) facilities. They were larger than the typical hospital (with an average of 346 beds, compared with 188 beds in non-participating hospitals) and likelier to be teaching hospitals.

2. Patient population characteristics

Beneficiaries who were treated in a hospital participating in Model 2 for a non-surgical cardiovascular episode were similar to all Medicare beneficiaries with the same MS-DRG. Both groups were similar with respect to gender, dual eligibility status, and Medicare eligibility due to disability (see Exhibit 69). The major difference between the BPCI and non-BPCI groups was related to age. The patients of BPCI providers were older than all beneficiaries with the same MS-DRG, with a larger proportion aged 80 and older (55.9% vs. 47.7%).

---

[53] The clinical episodes that compose the non-surgical cardiovascular group are: acute myocardial infarction; atherosclerosis; cardiac arrhythmia; chest pain; congestive heart failure; medical peripheral vascular disorders; and syncope & collapse.

Exhibit 69: Characteristics of BPCI Beneficiaries and All Medicare Beneficiaries with a Hospitalization for a Non-Surgical Cardiovascular MS-DRG, Model 2, Q4 2013-Q3 2014

Characteristics	Model 2 BPCI Beneficiaries with Non-Surgical Cardiovascular Episodes, Q4 2013 – Q3 2014 (N=6,387)		All Medicare Beneficiaries With Same MS-DRG Admission, Q3 2014 (N=187,913)	
	N	%	N	%
Age				
20-64	569	8.9%	21,797	11.6%
65-79	2,248	35.2%	76,450	40.7%
80+	3,570	55.9%	89,666	47.7%
Gender				
Female	3,493	54.7%	101,623	54.1%
Male	2,894	45.3%	86,290	45.9%
Medicaid and Disability				
Percent Eligible for Medicaid	1,502	23.5%	43,463	23.1%
Percent Disability, no ESRD	731	11.4%	22,323	11.9%

Source: Lewin analysis of Medicare claims and enrollment data for episodes that began Q4 2011 through Q3 2014 for BPCI providers and for all admissions for the same MS-DRG to non-participating hospitals in Q3 2014.

3.   *Change in Medicare standardized allowed payments*

There were no statistically significant changes in standardized allowed payment for Part A and B services during the anchor hospitalization and the 90-day PDP between BPCI and comparison group episodes. As seen in Exhibit 70, the total standardized allowed payment for Part A and B services during the anchor hospitalization and the 90-day PDP for beneficiaries treated by BPCI providers decreased from baseline to intervention periods, but this decrease was not statistically significant. However, for 30-day episodes with PAC use, there was a statistically significant decline in the total standardized allowed payment for patients treated by the four BPCI providers with 30-day episodes relative to the comparison group (-$1,625). We will continue to monitor the within bundle allowed payment amount in future quarters of the initiative. There were no statistically significant differences in the change for the allowed amount for any Part A or Part B service. There were also no statistically significant differences in the change in PAC use – both the first PAC setting and the intensity of PAC services (i.e. length of stay or number of home health visits) – relative to the episodes from matched comparison providers. There were no differences in case-mix measures between BPCI and comparison groups, which are monitored for indications that costs were shifted to episodes not under BPCI. Nor were there differences in spending before or after the episode definition, which could indicate that costs were shifted outside of the bundled period.

Exhibit 70: Impact of BPCI on Total Standardized Allowed Payment Inpatient Anchor Hospitalization and 90-day PDP for Non-Surgical Cardiovascular Episodes, Model 2, Q4 2011-Q3 2014

Measure	BPCI (N=6,274)		Comparison (N=6,264)		Diff-in-Diff estimate		
	Baseline	Intervention	Baseline	Intervention	Diff-in-Diff	LCI	UCI
Total Standardized Allowed Payment, Inpatient Anchor Hospitalization and 90-day PDP	$24,510	$24,383	$24,249	$23,962	$160	-$689	$1,009

* Denotes statistical significance at the 5% level.

LCI = lower confidence interval at the 5% level; UCI = upper confidence interval at the 5% level.

Note: Baseline is defined as episodes that began Q4 2011 through Q3 2012. Intervention is defined as episodes that began Q4 2013 through Q3 2014. Sample sizes reflect number of episodes initiated during the intervention period that met inclusion criteria for the outcome.

Source: Lewin analysis of Medicare claims and enrollment data for episodes that began Q4 2011 through Q3 2014 for BPCI and comparison providers.

4. Change in beneficiary quality of care

a. Claim and assessment-based measures

The quality of care for BPCI beneficiaries with non-surgical cardiovascular episodes was similar to that of beneficiaries of non-BPCI providers, as measured through various claims-based measures and PAC assessments. As seen in Exhibit 71, the change in hospital readmission rates, emergency department visits, and mortality following the anchor hospitalization for non-surgical cardiovascular episodes was not statistically different for beneficiaries of BPCI providers and comparison providers.   There were no statistically significant changes in assessment-based quality measures, from baseline to intervention period, between BPCI and comparison group beneficiaries with non-surgical cardiovascular episodes who received their first PAC treatment at a SNF, a HHA or an IRF.

Exhibit 71: Impact of BPCI on Claims-based Quality Measures for Non-Surgical Cardiovascular Episodes, Model 2, Q4 2011-Q3 2014

Measure	BPCI (N=6,316)		Comparison (N=6,325)		Diff-in-Diff estimate		
	Baseline	Intervention	Baseline	Intervention	Diff-in-Diff	LCI	UCI
Unplanned Readmission Rate, within 30 days	20.5%	19.9%	20.2%	19.1%	0.5	-1.4	2.3
Unplanned Readmission Rate, within 90 days	36.5%	36.7%	36.5%	35.3%	1.3	-1.0	3.5
Emergency Department Visit Rate, within 30 days	10.8%	11.4%	10.8%	11.9%	-0.4	-1.8	0.9
Emergency Department Visit Rate, within 90 days	21.8%	22.7%	21.6%	22.7%	-0.2	-2.1	1.8
Mortality, within 30 days	7.5%	7.3%	7.2%	7.5%	-0.5	-1.6	0.6

* Denotes statistical significance at the 5% level.

LCI = lower confidence interval at the 5% level; UCI = upper confidence interval at the 5% level.

Note: Baseline is defined as episodes that began Q4 2011 through Q3 2012. Intervention is defined as episodes that began Q4 2013 through Q3 2014. Sample sizes reflect number of episodes initiated during the intervention period that met inclusion criteria for the outcome.

Source: Lewin analysis of Medicare claims and enrollment data for episodes that began Q4 2011 through Q3 2014 for BPCI and comparison providers.

b. Beneficiary survey measures

Exhibits 72 and 73 below report the results for self-reported functional status improvement, functional status decline, and patient care experiences, respectively for beneficiaries with a non-surgical cardiovascular clinical episode. For beneficiaries with non-surgical cardiovascular episodes, there were no statistically significant differences in the proportion of BPCI and comparison respondents who improved or declined in any of the seven functional health measures.

Exhibit 72: Improvement and Decline in Functional Status for BPCI and Comparison Survey Respondents for Non-Surgical Cardiovascular Episodes, Model 2, October and November 2014

Functional Measure	BPCI Rate	Comparison Rate	Treatment Effect
Improvement in bathing, dressing, using the toilet, or eating	60.7% (347)	57.2% (342)	3.5 [-2.5, 9.4]
Improvement in walking without rest	27.4% (342)	27.2% (347)	0.2 [-5.6, 6.0]
Improvement in use of mobility device (i.e., less frequent)	37.8% (345)	40.2% (349)	-2.5 [-7.7, 2.8]
Improvement in using stairs	26.8% (335)	29.4% (333)	-2.5 [-8.5, 3.4]
Improvement in planning regular tasks	46.4% (345)	47.7% (353)	-1.3 [-7.4, 4.8]
Improvement in physical/emotional problems limiting social activities (i.e., less frequent)	45.3% (337)	51.0% (342)	-5.6 [-13.0, 1.6]
Improvement in pain limiting regular activities (i.e., less frequent)	45.4% (338)	46.6% (349)	-1.2 [-8.0, 5.6]
Decline in bathing, dressing, using the toilet, or eating	19.8% (344)	22.3% (340)	-2.5 [-8.6, 3.6]
Decline in walking without rest	37.4% (342)	37.0% (347)	0.4 [-5.6, 6.3]
Decline in use of mobility device (i.e., more frequent)	48.3% (354)	44.5% (373)	3.8 [-2.1,9.7]
Decline in using stairs	42.0% (335)	43.7% (333)	-1.7 [-7.8, 4.5]
Decline in planning regular tasks	27.5% (355)	30.1% (377)	-2.7 [-9.0, 3.7]
Decline in physical/emotional problems limiting social activities (i.e., more frequent)	27.8% (330)	27.5% (334)	0.3 [-6.4, 7.0]
Decline in pain limiting regular activities (i.e., more frequent)	20.8% (338)	21.7% (349)	-0.9 [-6.8, 5.0]

* Denotes statistical significance at the 5% level.

Notes: Treatment effect reported in percentage points. 95% confidence interval is in brackets. Survey results were weighted to adjust for sampling and nonresponse.

BPCI and comparison samples sizes for each question are reported in parentheses. The sample sizes vary between questions due to varying rates of item-response.

Source: Lewin analysis of Medicare beneficiary survey for Model 2 non-surgical cardiovascular BPCI and comparison group survey respondents.

There were no statistically significant differences between BPCI and comparison respondents in measures of care experiences, with the exception that BPCI respondents with non-surgical

cardiovascular episodes were significantly less likely to agree that they or their caregivers have been able to manage their health needs since returning home. 92.2 % of BPCI respondents indicated that they were able to manage their health needs, compared with 96.7% of comparison group respondents.

Exhibit 73: Health Care Experience for BPCI and Comparison Survey Respondents for Non-Surgical Cardiovascular Episodes, Model 2, October and November 2014

Health Care Experience Measures	BPCI Rate	Comparison Rate	Treatment Effect
Thinking about all the care you received in the hospital before and afterwards, how often did you, your family, or your caregiver get conflicting advice from medical staff about your treatment? 1[response = Never]	65.5% (341)	70.3% (352)	-4.8 [11.9, 2.3]
Thinking about all of the care you received in the hospital and afterwards, how often were the services you got appropriate for the level of care you needed? 1[response = Always]	55.9% (343)	63.1% (356)	-7.2 [-14.6, 0.2]
Thinking about when you left the hospital, were you discharged at the right time? 1[response = Yes]	86.3% (348)	86.0% (355)	0.3 [-5.0, 5.5]
Thinking about when you left the hospital listed in the cover letter, the medical staff took your preferences and those of your family or your caregiver into account in deciding what health care services you should have after you left the hospital. 1[response = Agree/Strongly Agree]	92.4% (307)	95.7% (308)	-3.7 [-7.3, 0.5]
Before you prepared to go home (or to someone else's home, or to an assisted living facility), you and your family or caregiver had a good understanding of how to take care of yourself. 1[response = Agree/Strongly Agree]	95.2% (297)	96.9% (315)	-1.7 [-5.0, 1.6]
Before you prepared to go home (or to someone else's home, or to an assisted living facility), medical staff clearly explained how to take your medications. 1[response = Agree/Strongly Agree]	95.8% (297)	93.7% (306)	2.1 [-1.8, 5.9]
Before you prepared to go home (or to someone else's home, or to an assisted living facility), medical staff clearly explained what follow-up appointments or treatments would be needed 1[response = Agree/Strongly Agree]	93.7% (293)	95.5% (310)	-1.8 [-5.5, 2.0]
Overall, since you returned home (or to someone else's home, or to an assisted living facility), you and your caregivers have been able to manage your health needs. 1[response = Agree/Strongly Agree]	92.2% (300)	96.7% (311)	-4.5* [-8.4, -0.6]
Overall, how satisfied are you with your recovery since you left the hospital? 1[response = Extremely/Quite a Bit]	61.4% (326)	66.7% (332)	-5.4 [-12.9, 2.1]

* Denotes statistical significance at the 5% level.

Notes: Results are unadjusted. Treatment effect reported in percentage points. 95% confidence interval is in brackets. Survey results were weighted to adjust for sampling and nonresponse.

BPCI and comparison samples sizes for each question are reported in parentheses. The sample sizes vary between questions due to varying rates of item-response.

Source: Lewin analysis of Medicare beneficiary survey for Model 2 non-surgical cardiovascular BPCI and comparison group survey respondents.

I.   Impact of BPCI on Non-Surgical Other Medical

1.   Participants

During the first four quarters of the initiative, 16 hospitals, or 15 % of hospital EIs in Model 2, participated in at least one non-surgical other medical episode.[54] There were 4,225 non-surgical other medical episodes initiated in Model 2-participating hospitals, accounting for 10% of all Model 2 episodes. Approximately 54% of these episodes were for sepsis MS-DRGs. All of the hospitals participating in BPCI with these clinical episodes were urban facilities and 94% were non-profits. They were larger than the typical hospital (with an average of 344 beds, compared with 188 beds in non-participating hospitals) and more likely to be teaching hospitals.

2.   Patient population characteristics

Beneficiaries who were treated in a hospital participating in Model 2 for a non-surgical other medical episode were similar to all Medicare beneficiaries treated in the same MS-DRG. Both groups were similar with respect to gender, dual eligibility status, and Medicare eligibility due to disability (see Exhibit 74). The major difference between BPCI beneficiaries and all Medicare beneficiaries was related to age. The patients of BPCI providers were older than all beneficiaries with the same MS-DRG, with a larger proportion aged 80 and older (48.5% vs. 40.1%) and a smaller proportion aged 20-64 (15.5 % vs. 20.4%).

Exhibit 74: Characteristics of BPCI Beneficiaries and All Medicare Beneficiaries with a Hospitalization for a Non-surgical Other Medical MS-DRG, Model 2, Q4 2013 - Q3 2014

Characteristics	Model 2 BCPI Beneficiaries With Non-Surgical Other Medical Episodes, Q4 2013 – Q3 2014 (N=4,039)		All Medicare Beneficiaries With Same MS-DRG Admission, Q3 2014 (N=335,937)	
	N	%	N	%
Age				
20-64	625	15.5%	68,453	20.4%
65-79	1,456	36.0%	132,901	39.6%
80+	1,958	48.5%	134,585	40.1%
Gender				
Female	2,290	56.7%	189,678	56.5%
Male	1,749	43.3%	146,259	43.5%
Other				
Percent Eligible for Medicaid	1,212	30.0%	108,361	32.3%
Percent Disability, no ESRD	715	17.7%	67,444	20.1%

Source: Lewin analysis of Medicare claims and enrollment data for episodes that began Q4 2011 through Q3 2014 for BPCI providers and for all admissions for the same MS-DRG to non-participating hospitals in Q3 2014.

---

[54] The clinical episodes that compose the non-surgical other medical group are: cellulitis; diabetes; nutritional and metabolic disorders; red blood cell disorders; renal failure; sepsis; and urinary tract infection.

3. Change in Medicare standardized allowed payments

There were no statistically significant differences in Medicare standardized allowed payments between BPCI and comparison group episodes. The average total standardized allowed payment for Parts A or B services during the anchor hospitalization and the 90-day PDP for beneficiaries treated by BPCI providers declined from baseline to intervention periods, but this reduction was not statistically significant relative to the change for beneficiaries treated by comparison providers (see Exhibit 75). There were also no statistically significant differences in the change for the allowed amount for any Part A or Part B service (i.e. SNF, IRF, imaging and lab). Relative to the episodes from comparison providers, there were no statistically significant differences in the change in PAC use – both the first PAC setting and the intensity of PAC services (i.e. length of stay or number of home health visits). There were no indications that these results were achieved by shifting services outside the bundle period or not covered in the bundle definition.

Exhibit 75: Impact of BPCI on Total Standardized Allowed Payment Inpatient Anchor Hospitalization and 90-day PDP for Non-surgical Other Medical Episodes, Model 2, Q4 2011-Q3 2014

Measure	BPCI (N=3,944)		Comparison (N=3,933)		Diff-in-Diff estimate		
	Baseline	Intervention	Baseline	Intervention	Diff-in-Diff	LCI	UCI
Total Standardized Allowed Payment, Inpatient Anchor Hospitalization and 90-day PDP	$30,649	$29,728	$28,374	$27,548	-$96	-$1,695	$1,504

* Denotes statistical significance at the 5% level.

LCI = lower confidence interval at the 5% level; UCI = upper confidence interval at the 5% level.

Note: Baseline is defined as episodes that began Q4 2011 through Q3 2012. Intervention is defined as episodes that began Q4 2013 through Q3 2014. Sample sizes reflect number of episodes initiated during the intervention period that met inclusion criteria for the outcome.

Source: Lewin analysis of Medicare claims and enrollment data for episodes that began Q4 2011 through Q3 2014 for BPCI providers and a comparison group.

4. Change in beneficiary quality of care

The quality of care for BPCI beneficiaries with non-surgical other medical episodes was similar to that of beneficiaries of non-BPCI providers, as measured through various claim-based measures and PAC assessments. The change in hospital readmission rates, emergency department visits, and mortality following the anchor hospitalization for non-surgical other medical episodes was not statistically different for patients of BPCI providers and comparison providers (see Exhibit 76). There were no statistically significant changes in assessment-based quality measures, from baseline to intervention period, between BPCI and comparison group beneficiaries with non-surgical other medical episodes who received their first PAC treatment at a SNF, a HHA or an IRF.

Exhibit 76: Impact of BPCI on Claims-based Quality Measures for Non-Surgical Other Medical Episodes, Model 2, Q4 2011-Q3 2014

Measure	BPCI (N=3,992)		Comparison (N=4,010)		Diff-in-Diff estimate		
	Baseline	Intervention	Baseline	Intervention	Diff-in-Diff	LCI	UCI
Unplanned Readmission Rate, within 30 days	16.5%	16.3%	17.2%	16.3%	0.7	-1.3	2.7
Unplanned Readmission Rate, within 90 days	29.9%	29.2%	30.9%	29.7%	0.5	-1.6	2.6
Emergency Department Visit Rate, within 30 days	9.7%	10.1%	9.4%	9.4%	0.4	-1.0	1.8
Emergency Department Visit Rate, within 90 days	18.8%	21.1%	18.9%	19.6%	1.6	-0.6	3.9
Mortality, within 30 days	11.2%	9.8%	9.3%	9.4%	-1.5	-3.2	0.2

* Denotes statistical significance at the 5% level.

LCI = lower confidence interval at the 5% level; UCI = upper confidence interval at the 5% level.

Note: Baseline is defined as episodes that began Q4 2011 through Q3 2012. Intervention is defined as episodes that began Q4 2013 through Q3 2014. Sample sizes reflect number of episodes initiated during the intervention period that met inclusion criteria for the outcome.

Source: Lewin analysis of Medicare claims and enrollment data for episodes that began Q4 2011 through Q3 2014 for BPCI providers and a comparison group.

J. Impact of BPCI on Non-Surgical Neurovascular

1. Participants

During the first year of the initiative, 15 hospitals, or 14% of hospital EIs in Model 2, participated in at least one non-surgical neurovascular episode.[55] There were 1,109 non-surgical neurovascular episodes initiated in Model 2-participating hospitals, accounting for nearly 3% of all Model 2 episodes. The vast majority (91%) of these episodes were for stroke MS-DRGs. All of the hospitals participating in BPCI with these clinical episodes were urban facilities and were non-profit organizations. They were larger than the typical hospital (with an average of 370 beds, compared with 188 beds in non-participating hospitals) and likelier to be teaching hospitals.

2. Patient population characteristics

Beneficiaries who were treated in a hospital participating in Model 2 for a non-surgical neurovascular episode were similar to all Medicare beneficiaries treated in the same MS-DRG with respect to gender and dual eligibility status (see Exhibit 77). The major differences between BPCI beneficiaries and all Medicare beneficiaries were related to age and Medicare eligibility due to disability. The patients of BPCI providers were older than all beneficiaries with the same MS-DRG, with a larger proportion aged 80 and older (53.7% vs. 45.3%) and a smaller proportion aged 20-64 (6.1% vs. 11.7%). Among BPCI beneficiaries, 7.8% were eligible for Medicare due to disability, compared to 13.1% of all Medicare beneficiaries.

---

[55] The clinical episodes that compose the non-surgical neurovascular group are stroke and transient ischemia.

Exhibit 77: Characteristics of BPCI Beneficiaries and All Medicare Beneficiaries with a Hospitalization for a Non-Surgical Neurovascular MS-DRG, Model 2, Q4 2013 - Q3 2014

Characteristics	Model 2 BPCI Beneficiaries with Non-Surgical Neurovascular Episodes, Q4 2013 – Q3 2014 (N=1,057)		All Medicare Beneficiaries With Same MS-DRG Admission, Q3 2014 (N=15,065)	
	N	%	N	%
Age				
20-64	64	6.1%	1,756	11.7%
65-79	425	40.2%	6,482	43.0%
80+	568	53.7%	6,827	45.3%
Gender				
Female	602	57.0%	8,915	59.2%
Male	455	43.0%	6,150	40.8%
Medicaid and Disability				
Percent Eligible for Medicaid	217	20.5%	3,196	21.2%
Percent Disability, no ESRD	82	7.8%	1,971	13.1%

Source: Lewin analysis of Medicare claims and enrollment data for episodes that began Q4 2011 through Q3 2014 for BPCI providers and for all admissions for the same MS-DRG to non-participating hospitals in Q3 2014.

3. *Change in Medicare standardized allowed payment amounts*

There were no statistically significant changes in standardized allowed payments between BPCI and comparison group episodes. The average total standardized allowed payment for Parts A or B services during the anchor hospitalization and the 90-day PDP for beneficiaries treated by BPCI providers declined from baseline to intervention periods, but this reduction was not statistically significant relative to the change for beneficiaries treated by matched comparison providers (see Exhibit 78). Furthermore, there were no statistically significant differences in the change for the allowed amount for any Part A or Part B service (i.e. SNF, IRF, imaging and lab). There were also no statistically significant differences in the change in PAC use – both the first PAC setting and the intensity of PAC services (i.e. length of stay or number of home health visits) - relative to the episodes from matched comparison providers. There were no indications that these results were achieved by shifting services outside the bundle period or not covered in the bundle definition.

Exhibit 78: Impact of BPCI on Total Standardized Allowed Payment Inpatient Anchor Hospitalization and 90-day PDP for Non-Surgical Neurovascular Episodes, Model 2, Q4 2011-Q3 2014

Measure	BPCI (N=1,039)		Comparison (N=1,038)		Diff-in-Diff estimate		
	Baseline	Intervention	Baseline	Intervention	Diff-in-Diff	LCI	UCI
Total Standardized Allowed Payment, Inpatient Anchor Hospitalization and 90-day PDP	$32,022	$30,787	$31,375	$30,333	-$194	-$2,987	$2,600

* Denotes statistical significance at the 5% level.

LCI = lower confidence interval at the 5% level; UCI = upper confidence interval at the 5% level.

Note: Baseline is defined as episodes that began Q4 2011 through Q3 2012. Intervention is defined as episodes that began Q4 2013 through Q3 2014. Sample sizes reflect number of episodes initiated during the intervention period that met inclusion criteria for the outcome.

Source: Lewin analysis of Medicare claims and enrollment data for episodes that began Q4 2011 through Q3 2014 for BPCI providers and a comparison group.

4. Change in beneficiary quality of care

The quality of care for BPCI patients with non-surgical neurovascular episodes was similar to that of patients of BPCI comparison providers, as measured through various claim-based measures and PAC assessments. As seen in Exhibit 79, the change in hospital readmission rates, emergency department visits, and mortality following the anchor hospitalization for non-surgical neurovascular episodes was not statistically different for patients of BPCI providers and comparison providers.   There were no statistically significant changes in assessment-based quality measures from baseline to intervention period between BPCI and comparison group beneficiaries with non-surgical neurovascular episodes who received their first PAC treatment at a SNF, a HHA or an IRF.

Exhibit 79: Impact of BPCI on Claims-based Quality Measures for Non-Surgical Neurovascular Episodes, Model 2, Q4 2011-Q3 2014

Measure	BPCI (N=1,050)		Comparison (N=1,051)		Diff-in-Diff estimate		
	Baseline	Intervention	Baseline	Intervention	Diff-in-Diff	LCI	UCI
Unplanned Readmission Rate, within 30 days	12.7%	11.5%	13.5%	11.9%	0.5	-2.5	3.5
Unplanned Readmission Rate, within 90 days	23.4%	22.0%	22.6%	22.3%	-1.0	-6.3	4.2
Emergency Department Visit Rate, within 30 days	8.2%	9.2%	8.4%	8.6%	0.7	-2.4	3.8
Emergency Department Visit Rate, within 90 days	18.4%	18.4%	18.0%	18.0%	0.1	-4.6	4.7
Mortality, within 30 days	9.9%	10.3%	9.7%	11.8%	-1.7	-5.6	2.2

* Denotes statistical significance at the 5% level.

LCI = lower confidence interval at the 5% level; UCI = upper confidence interval at the 5% level.

Note: Baseline is defined as episodes that began Q4 2011 through Q3 2012. Intervention is defined as episodes that began Q4 2013 through Q3 2014. Sample sizes reflect number of episodes initiated during the intervention period that met inclusion criteria for the outcome.

Source: Lewin analysis of Medicare claims and enrollment data for episodes that began Q4 2011 through Q3 2014 for BPCI and comparison providers.

K. Impact of BPCI on Non-Surgical Respiratory

1. Participants

During the first year of the initiative, 33 hospitals, or 30% of EIs in Model 2, participated in at least one non-surgical respiratory episode.[56] There were 5,805 non-surgical respiratory episodes initiated in Model 2-participating hospitals, accounting for 14% of all Model 2 episodes. Approximately 49% of these episodes were for simple pneumonia and respiratory infections, and 45% were for chronic obstructive pulmonary disease, bronchitis, or asthma MS-DRGs. Nearly all (94%) of the hospitals participating in BPCI with these clinical episodes were urban facilities and nearly all (94%) were non-profits. They were larger than the typical hospital (with an average of 344 beds, compared with 188 beds in non-participating hospitals) and likelier to be teaching hospitals.

2. Patient population characteristics

Beneficiaries who were treated in a hospital participating in Model 2 for a non-surgical respiratory episode were similar to all Medicare beneficiaries treated in the same MS-DRG. Both groups were similar with respect to gender, dual eligibility status, and Medicare eligibility due to disability (see Exhibit 80). The major difference between the BPCI and non-BPCI groups was related to age. The patients of BPCI providers were older than all beneficiaries with the same MS-DRG, with a larger proportion aged 80 and older (42.9% vs. 36.1%).

Exhibit 80: Characteristics of BPCI Beneficiaries and All Medicare Beneficiaries with a Hospitalization for Non-Surgical Respiratory MS-DRG, Model 2, Q4 2013-Q3 2014

Characteristics	Model 2 BPCI Beneficiaries with Non-Surgical Respiratory Episodes, Q4 2013 – Q3 2014 (N=5,532)		All Medicare Beneficiaries With Same MS-DRG Admission, Q3 2014 (N=201,913)	
	N	%	N	%
Age				
20-64	907	16.4%	41,713	20.7%
65-79	2,253	40.7%	87,404	43.3%
80+	2,372	42.9%	72,796	36.1%
Gender				
Female	3,185	57.6%	111,281	55.1%
Male	2,347	42.4%	90,632	44.9%
Medicaid and Disability				
Percent Eligible for Medicaid	1,617	29.2%	66,017	32.7%
Percent Disability, no ESRD	1,092	19.7%	43,763	21.7%

Source: Lewin analysis of Medicare claims and enrollment data for episodes that began Q4 2011 through Q3 2014 for BPCI providers and for all admissions for the same MS-DRG to non-participating hospitals in Q3 2014.

---

[56] The clinical episodes that compose the non-surgical respiratory group are: chronic obstructive pulmonary disease, bronchitis, and asthma; simple pneumonia and respiratory infections; and other respiratory.

3. Change in Medicare standardized allowed payments

There were no statistically significant differences between BPCI non-surgical respiratory episodes and episodes in comparison providers in Medicare standardized allowed payments. The average total standardized allowed payment for Parts A or B services during the anchor hospitalization and the 90-day PDP for patients treated by BPCI providers declined from baseline to intervention periods, but this reduction was not statistically significant relative to the change for beneficiaries treated by comparison providers (see Exhibit 81). Furthermore, there were no statistically significant differences in the allowed amount for any Part A or Part B service (i.e. SNF, IRF, imaging and lab) or PAC use – both the first PAC setting and the intensity of PAC services (i.e. length of stay or number of home health visits) - relative to the episodes from comparison providers. There were no indications that BPCI participating providers shifted services outside of the bundle definition or bundle period.

Exhibit 81: Impact of BPCI on Total Standardized Allowed Payment Inpatient Anchor Hospitalization and 90-day PDP for Non-Surgical Respiratory Episodes, Model 2, Q4 2011-Q3 2014

Measure	BPCI (N=5,409)		Comparison (N=5,419)		Diff-in-Diff estimate		
	Baseline	Intervention	Baseline	Intervention	Diff-in-Diff	LCI	UCI
Total Standardized Allowed Payment, Inpatient Anchor Hospitalization and 90-day PDP	$22,769	$22,011	$22,048	$21,322	-$32	-$881	$817

* Denotes statistical significance at the 5% level.

LCI = lower confidence interval at the 5% level; UCI = upper confidence interval at the 5% level.

Note: Baseline is defined as episodes that began Q4 2011 through Q3 2012. Intervention is defined as episodes that began Q4 2013 through Q3 2014. Sample sizes reflect number of episodes initiated during the intervention period that met inclusion criteria for the outcome.

Source: Lewin analysis of Medicare claims and enrollment data for episodes that began Q4 2011 through Q3 2014 for BPCI and comparison providers.

4. Change in beneficiary quality of care

a. Claim and assessment-based measures

The quality of care for BPCI beneficiaries with non-surgical respiratory episodes was similar to that of patients of non-BPCI providers, as measured through various claim-based measures and PAC assessments. As seen in Exhibit 82, the change in hospital readmission rates, emergency department visits, and mortality following the anchor hospitalization for non-surgical respiratory episodes was not statistically different for patients of BPCI providers and comparison providers. There were no statistically significant changes in assessment-based quality measures from baseline to intervention period between BPCI and comparison group beneficiaries and non-surgical respiratory episodes who received their first PAC treatment at a SNF, a HHA or an IRF.

Exhibit 82: Impact of BPCI on Claims-based Quality Measures for Non-Surgical Respiratory Episodes, Model 2, Q4 2011-Q3 2014

Measure	BPCI (N=5,492)		Comparison (N=5,478)		Diff-in-Diff estimate		
	Baseline	Intervention	Baseline	Intervention	Diff-in-Diff	LCI	UCI
Unplanned Readmission Rate, within 30 days	17.1%	16.0%	16.8%	15.4%	0.3	-1.3	1.9
Unplanned Readmission Rate, within 90 days	32.0%	30.1%	31.2%	29.2%	0.1	-1.9	2.1
Emergency Department Visit Rate, within 30 days	10.9%	11.0%	10.9%	11.1%	-0.1	-1.7	1.5
Emergency Department Visit Rate, within 90 days	22.4%	24.2%	21.4%	22.7%	0.5	-1.8	2.7
Mortality, within 30 days	6.4%	6.5%	6.4%	6.9%	-0.4	-1.4	0.7

* Denotes statistical significance at the 5% level.

LCI = lower confidence interval at the 5% level; UCI = upper confidence interval at the 5% level.

Note: Baseline is defined as episodes that began Q4 2011 through Q3 2012. Intervention is defined as episodes that began Q4 2013 through Q3 2014. Sample sizes reflect number of episodes initiated during the intervention period that met inclusion criteria for the outcome.

Source: Lewin analysis of Medicare claims and enrollment data for episodes that began Q4 2011 through Q3 2014 for BPCI and comparison providers.

b.   Beneficiary survey measures

Beneficiaries with non-surgical respiratory episodes comprised one of three Model 2 strata in Wave 2. BPCI respondents with non-surgical respiratory episodes were 7.1 percentage points less likely than similar comparison respondents to report a decline in their ability to walk without rest subsequent to treatment (Exhibit 83).

Exhibit 83: Improvement and Decline in Functional Status for BPCI and Comparison Survey Respondents with Non-Surgical Respiratory Episodes, Model 2, October and November 2014

Functional Measure	BPCI Rate	Comparison Rate	Treatment Effect
Improvement in bathing, dressing, using the toilet, or eating	60.3% (266)	59.7% (271)	0.6 [-6.2, 7.4]
Improvement in walking without rest	30.8% (262)	29.2% (269)	1.6 [-5.1, 8.2]
Improvement in use of mobility device (i.e., less frequent)	41.7% (268)	44.2% (270)	-2.5 [-7.4, 2.4]
Improvement in using stairs	28.2% (261)	24.4% (265)	3.9 [-3.2, 11.0]
Improvement in planning regular tasks	44.8% (267)	44.4% (274)	0.5 [-6.8, 7.7]
Improvement in physical/emotional problems limiting social activities (i.e., less frequent)	45.1% (271)	45.8% (273)	-0.7 [-8.6, 7.2]
Improvement in pain limiting regular activities (i.e., less frequent)	44.0% (266)	42.7% (272)	1.3 [-6.6, 9.2]

Functional Measure	BPCI Rate	Comparison Rate	Treatment Effect
Decline in bathing, dressing, using the toilet, or eating	18.1% (266)	21.1% (271)	-3.0 [-9.4, 3.4]
Decline in walking without rest	27.9% (262)	34.9% (269)	-7.1* [-13.5, -0.6]
Decline in use of mobility device (i.e., more frequent)	42.5% (268)	40.1% (270)	2.4 [-3.5, 8.2]
Decline in using stairs	38.5% (261)	42.9% (265)	-4.4 [-10.4, 1.6]
Decline in planning regular tasks	25.6% (267)	25.7% (274)	-0.1 [-6.6, 6.3]
Decline in physical/emotional problems limiting social activities (i.e., more frequent)	27.9% (271)	30.5% (273)	-2.5 [-9.8, 4.7]
Decline in pain limiting regular activities (i.e., more frequent)	24.2% (266)	24.6% (272)	-0.5 [-7.2, 6.3]

* Denotes statistical significance at the 5% level.

Notes: Treatment effect reported in percentage points. 95% confidence interval is in brackets. Survey results were weighted to adjust for sampling and nonresponse

BPCI and comparison samples sizes for each question are reported in parentheses. The sample sizes vary between questions due to varying rates of item-response.

Source: Lewin analysis of Medicare beneficiary survey for Model 2 non-surgical cardiovascular BPCI and comparison group survey respondents.

There were no other statistically significant differences in rates of improvement or decline in functional status between BPCI and comparison respondents with non-surgical respiratory episodes. There were no statistically significant differences for beneficiary perception of care experience, including overall satisfaction with recovery, on any measure between BPCI and comparison respondents with non-surgical respiratory episodes (Exhibit 84).

Exhibit 84: Health Care Experience for BPCI and Comparison Survey Respondents with Non-surgical Respiratory Episodes, Model 2, October and November 2014

Health Care Experience Measures	BPCI Rate	Comparison Rate	Treatment Effect
Thinking about all the care you received in the hospital before and afterwards, how often did you, your family, or your caregiver get conflicting advice from medical staff about your treatment? 1[response = Never]	65.3% (274)	69.0% (274)	-3.7 [11.8, 4.4]
Thinking about all of the care you received in the hospital and afterwards, how often were the services you got appropriate for the level of care you needed? 1[response = Always]	55.6% (273)	58.5% (271)	-2.9 [-11.5, 5.7]
Thinking about when you left the hospital, were you discharged at the right time? 1[response = Yes]	86.7% (274)	83.0% (272)	3.7 [-2.8, 10.1]

Health Care Experience Measures	BPCI Rate	Comparison Rate	Treatment Effect
Thinking about when you left the hospital listed in the cover letter, the medical staff took your preferences and those of your family or your caregiver into account in deciding what health care services you should have after you left the hospital. 1[response = Agree/Strongly Agree]	92.1% (226)	90.8% (236)	1.3 [3.7, 6.2]
Before you prepared to go home (or to someone else's home, or to an assisted living facility), you and your family or caregiver had a good understanding of how to take care of yourself. 1[response = Agree/Strongly Agree]	95.3% (238)	94.7% (227)	0.6 [-3.6, 4.8]
Before you prepared to go home (or to someone else's home, or to an assisted living facility), medical staff clearly explained how to take your medications. 1[response = Agree/Strongly Agree]	96.5% (236)	93.9% (222)	2.7 [-1.3, 6.6]
Before you prepared to go home (or to someone else's home, or to an assisted living facility), medical staff clearly explained what follow-up appointments or treatments would be needed 1[response = Agree/Strongly Agree]	95.4% (225)	97.4% (226)	-2.0 [-5.4, 1.3]
Overall, since you returned home (or to someone else's home, or to an assisted living facility), you and your caregivers have been able to manage your health needs. 1[response = Agree/Strongly Agree]	92.8% (230)	94.6% (230)	-1.8 [-6.7, 3.0]
Overall, how satisfied are you with your recovery since you left the hospital? 1[response = Extremely/Quite a Bit]	61.9% (259)	63.6% (257)	-1.8 [10.2, 6.7]

* Denotes statistical significance at the 5% level.

Notes: Results are unadjusted. Treatment effect reported in percentage points. 95% confidence interval is in brackets.

BPCI and comparison samples sizes for each question are reported in parentheses. The sample sizes vary between questions due to varying rates of item-response.

Source: Lewin analysis of Medicare beneficiary survey for BPCI and comparison group survey respondents.

IV. Model 3 Results

This section presents information about the experience of Model 3 BPCI participants and their episodes of care, organized by research question. Similar to Model 2, quantitative analyses of Medicare claims data are based on episodes initiated during the first full year of BPCI (Q4 2013 to Q3 2014); patient assessment data are based on episodes initiated during the first three quarters of BPCI (Q4 2013 to Q2 2014); Awardee-submitted data are based on calendar year 2014 or Q4 2014 depending on the measure; and qualitative data from interviews reflect the first six quarters of the BPCI initiative (Q4 2013 to Q1 2015) while site visits results reflect the first seven quarters of the BPCI initiative (Q4 2013 to Q2 2015). Survey data are based on episodes initiated during May/June 2014 and October/November 2014. The quantitative outcomes are risk-adjusted as described in Section II.D.2 above and as noted, because of the small sample sizes, comparison providers were matched on only a small number of characteristics.

A. Characteristics of the Initiative and Participants

1. Participants

This section describes participants that joined BPCI through Q3 2014. Twenty active Awardees with 94 EIs were in Model 3 during the first year of the initiative, including 63 skilled nursing facilities (SNFs), 28 home health agencies (HHAs), 1 inpatient rehabilitation facility (IRF), 1 long-term care hospital (LTCH), and 1 physician group practice (PGP). In the following exhibits, BPCI SNF and BPCI HHA EIs are compared to all non-participating SNFs and HHAs.[57] See **Appendix Q** for a detailed description of the growth of the initiative through July 2015.

Exhibit 85 compares the first year BPCI-participating SNF EIs to non-participating SNFs. A higher proportion of BPCI-participating SNF EIs were for-profit organizations (83%) compared with non-participating SNFs (69%). No participating SNFs were government facilities compared with 5% of non-participating SNFs. An average of 17% of SNF participants were non-profits, compared to 26% of non-participating SNFs. Participating SNFs were less likely to be part of a chain than non-participating SNFs (17% vs. 23%).

Geographically, all BPCI-participating SNFs were located in urban areas, compared with only 69% of non-participating SNFs. An average of 56% of participating SNFs had an IRF in their CBSA, compared with only 29% of non-participants. The presence of an IRF suggests the availability of additional options for post-acute care in participating SNF markets. The greater presence of IRFs in participating SNF markets may be related to their location in urban settings, which likely have more post-acute care options than rural settings. Participating SNFs had higher bed counts on average than non-participating SNFs (148 vs. 110).

Although BPCI-participating SNFs were more likely for-profit, urban, non-chain organizations, they share similarities with non-participating SNFs on their stand-alone setting, capacity, market share, and quality ranking. Only very small proportions of participating SNFs (3%) and non-

---

[57] Due to the small number of IRFs, LTCHs, and PGPs, we do not compare them to the universe of IRFs, LTCHs, and PGPs in this report.

participating SNFs (5%) were based in hospitals. The average SNF market share was similar for both groups (under 10%), as was the Nursing Home Compare overall score (3.6 vs. 3.3).

BPCI-participating SNF EI episodes had higher standardized payments for the 90 day PDP in 2011 than non-participating SNF episodes. The difference varied by clinical episode group; the standardized payments were 7% higher for orthopedic surgery episodes in BPCI-participating SNFs and 22% higher for non-surgical cardiovascular episodes.

Exhibit 85: Characteristics of BPCI-participating SNF EIs and Non-participating SNFs, Model 3, Q4 2013 – Q3 2014

	BPCI SNF EIs (N=63)		Non-participating SNFs (N=13,287)	
	N	**%**	**N**	**%**
Ownership				
For Profit	52	83%	9,168	69%
Government	0%	0%	664	5%
Non-Profit	11	17%	3,455	26%
Urban/Rural				
Rural	0	0%	4,119	31%
Urban	63	100%	9,168	69%
IRF in CBSA				
Yes	35	56%	3,853	29%
Hospital-Based				
Yes	2	3%	664	5%
Part of Chain				
Yes	11	17%	3,056	23%
	Mean		**Mean**	
Bed Count	148		110	
Number of BPCI Episodes	318		110	
SNF Market Share	3%		6%	
Nursing Home Overall Score	3.6		3.3	
Standardized Part A Allowed Payment 90 day PDP, 2011				
Clinical Episode Group (N, BPCI admissions 2011)	**Mean**		**Mean**	
Orthopedic surgery (3795)	$14,243		$13,255	
Non-surgical respiratory (2008)	$19,607		$16,287	
Non-surgical cardiac (2489)	$19,691		$16,131	

Source: Lewin analysis of 2013 Provider of Service (POS) and 2011 Medicare claims. BPCI participating SNFs are defined as Episode Initiators, Q4 2013 – Q3 2014. Non-Participating SNFs are all other SNFs.

Exhibit 86 describes the Model 3 BPCI-participating HHAs in comparison to all non-participating HHAs. Small numbers of participating HHAs (28 EIs) may drive the variation described below.

BPCI-participating HHAs differed from non-participating HHAs on organization type and staff capacity. No BPCI-participating HHAs were government organizations compared with 5% of non-participating HHAs. A higher proportion of participating HHAs were for-profit (93%) compared with 79% of non-participants. In terms of organizational capacity, BPCI-participating HHAs had more employed nurses on average than did non-participating HHAs (70 vs. 9), suggesting that participating HHAs tended to be larger. Of note, the largest HHA in the nation participates in the non-surgical cardiovascular group, which inflates the average. After removing this HHA, the average number of employed nurses in HHAs participating in non-surgical cardiovascular episodes dropped to 15. The majority of both BPCI-participating and non-participating HHAs were in urban areas (79% and 83%, respectively). BPCI-participating HHA EI episodes had higher standardized payments for the 90 day PDP in 2011 relative to episodes in non-participating HHAs. The difference in standardized payments varied by clinical episode group; the standardized payments were 1% higher for non-surgical cardiovascular episodes in BPCI-participating HHAs and 13% higher for non-surgical respiratory episodes.

Exhibit 86: Characteristics of BPCI-participating HHA EIs and Non-participating HHAs, Model 3, Q4 2013 – Q3 2014

	Non-BPCI HHA EIs (N=28)		Non-participating HHAs (N=12,226)	
	N	%	N	%
Ownership				
For Profit	26	93%	9,659	79%
Government	0	0%	611	5%
Non-Profit	2	7%	1,834	15%
Urban/Rural				
Rural	6	21%	2,078	17%
Urban	22	79%	10,148	83%
Part of Chain				
Yes	24	86%	3,423	28%
	Mean		**Mean**	
Number of Employed Nurses in HHA	71		9	
Number of BPCI Episodes	1,729		113	
Standardized Part A Allowed Payment 90 day PDP, 2011				
Clinical Episode Group (N, BPCI admissions 2011)	**Mean**		**Mean**	
Non-surgical respiratory (1546)	$11,288		$9,803	
Non-surgical cardiac (5465)	$9,673		$9,607	

Source: Lewin analysis of 2013 POS and 2011 Medicare claims. BPCI participating HHAs are defined as Episode Initiators, Q4 2013 – Q3 2014. Non-Participant HHAs are all other HHAs.

a.   *Awardee submitted baseline characteristics*

Model 3 Awardees and EIs submitted information about their experiences that may have contributed to their implementation of BPCI through an on-line data reporting tool beginning in Q1 2015. This included patient payer mix, prior experience with care redesign and payment

initiatives, and medication reconciliation. We collected data representing the Awardee and EI experience during various time periods, including prior to joining BPCI and during 2014.

Patients by payer type

Among Model 3 EIs, we calculated patient mix for 62 SNFs and 28 HHAs, as illustrated in Exhibit 87.[58] Payer mix was based on the reported number of patients in 2014 with a primary payer of Medicaid, Medicare, private, or other. Medicare patients comprised the highest portion of total patients across SNFs and HHAs, with averages of 45.8% and 80.2%, respectively.

Exhibit 87: Patient Payer Mix, Model 3 SNFs and HHAs, 2014

	Model 3 SNF EIs N=62				Model 3 HHA EIs N=28			
	Mean	**Min**	**Median**	**Max**	**Mean**	**Min**	**Median**	**Max**
Percentage of Medicaid Patients	12.9%	0.0%	7.9%	68.5%	1.5%	0.0%	0.0%	19.5%
Percentage of Medicare Patients	45.8%	7.1%	48.0%	92.0%	80.2%	62.1%	76.0%	99.9%
Percentage of Private Payer Patients	6.5%	0.2%	2.3%	45.4%	18.1%	0.1%	22.9%	33.8%
Percentage of Other Payer Patients	34.8%	0.3%	36.3%	70.1%	0.2%	0.0%	0.0%	4.9%

Source: Lewin analysis of Awardee-submitted data collected February 2015 for Model 3 EIs participating in BPCI between Q4 2013 – Q4 2014.

Prior experience

Exhibit 88 summarizes the prior participation of Model 3 SNFs and HHAs in care redesign initiatives and payment incentives.[59] The majority of SNFs reported having experience in redesign of care pathways (77.4%); however, a much smaller percentage of EIs reported having prior experience in the remaining four formal categories of care redesign, ranging from 11.3% for system changes to support care to 22.6% for care coordination. HHAs, on the other hand, reported broad experience in care redesign initiatives, with over 60% of HHAs reporting prior experience in each of the five structured categories.

The most common type of prior payment incentives experience among Model 3 EIs was pay for performance, with 22.6% of SNFs and 53.6% of HHAs reporting prior participation. Neither SNFs nor HHAs reported prior experience with bundled payment initiatives.

---

[58] We also collected data from 1 LTCH, 1 IRF, and 1 Model 3 PGP. However, due to limited sample sizes, we did not calculate payer mix for these facilities.

[59] We also collected data from 1 LTCH, 1 IRF, and 1 Model 3 PGP. However, due to limited sample sizes, we did not present care redesign and payment incentives for these facilities.

Exhibit 88: Care Redesign and Payment Incentives Experience, Model 3 SNFs and HHAs, Prior to BPCI Participation

	Model 3 SNF EIs N=62		Model 3 HHA EIs N=28	
	N	%	N	%
Prior experience in care redesign initiatives:				
Redesign of Care Pathways	48	77.4%	17	60.7%
Enhancements in Care Delivery	8	12.9%	17	60.7%
Patient Activation, Engagement & Risk Management	11	17.7%	17	60.7%
Care Coordination	14	22.6%	18	64.3%
System Changes to Support Care	7	11.3%	17	60.7%
Other Redesign Activities	2	3.2%	0	0.0%
Prior experience in payment incentives:				
Bundled Payments	0	0.0%	0	0.0%
Pay for Performance	14	22.6%	15	53.6%
Shared Savings	1	1.6%	0	0.0%
Other Payment Incentives	4	6.5%	1	3.6%

Source: Lewin analysis of Awardee-submitted data collected February 2015 for Model 3 EIs participating in BPCI between Q4 2013 – Q4 2014.

b. *Awardee submitted quality measures*

Medication reconciliation at admission

Model 3 EIs reported the number of BPCI patient admissions among those 65 years of age and older for whom medications were reconciled at admission and the total number of admissions among this population during Q4 2014. As displayed in Exhibit 89, on average, medication reconciliation occurred during 95% of SNF admissions and 100% of HHA admissions. The majority of Model 3 EIs reported that medication reconciliation occurred upon every admission.

Exhibit 89: Proportion of BPCI Beneficiaries receiving Medication Reconciliation at Admission to PAC Facility, Model 3, Q4 2014

Model 3 EI type	N	Proportion of BPCI Beneficiaries receiving Medication Reconciliation at PAC admission			
		Mean	Median	Minimum	Maximum
SNF	62	95%	100%	0%	100%
HHA	27	100%	100%	96%	100%
IRF	1	100%	100%	100%	100%
LTCH	1	100%	100%	100%	100%
PGP	1	53%	53%	53%	53%

Source: Lewin analysis of Awardee-submitted data collected February 2015 for Model 3 EIs participating in BPCI during Q4 2014.

2. Market characteristics

The 94 Model 3 BPCI EIs that participated in the first four quarters of the BPCI initiative were located throughout the country. Exhibits 90 to 94 display the geographic locations of the episode-initiating SNF, IRF, HHA, LTCH, and PGP-affiliated facilities. In this section, we compare characteristics of the markets where Model 3 BPCI PAC providers were located with the areas where there were no BPCI-participating PAC providers during the first year of the initiative.[60] It should be noted that the non-BPCI markets include those markets with EIs in Phase 1 of the program. Thus, some non-BPCI markets will become BPCI markets with the large influx of providers who transition to Phase 2 in Q2 and Q3 2015.

Exhibit 90: Number of BPCI Participating SNFs by CBSA, Model 3, Q4 2013 – Q3 2014
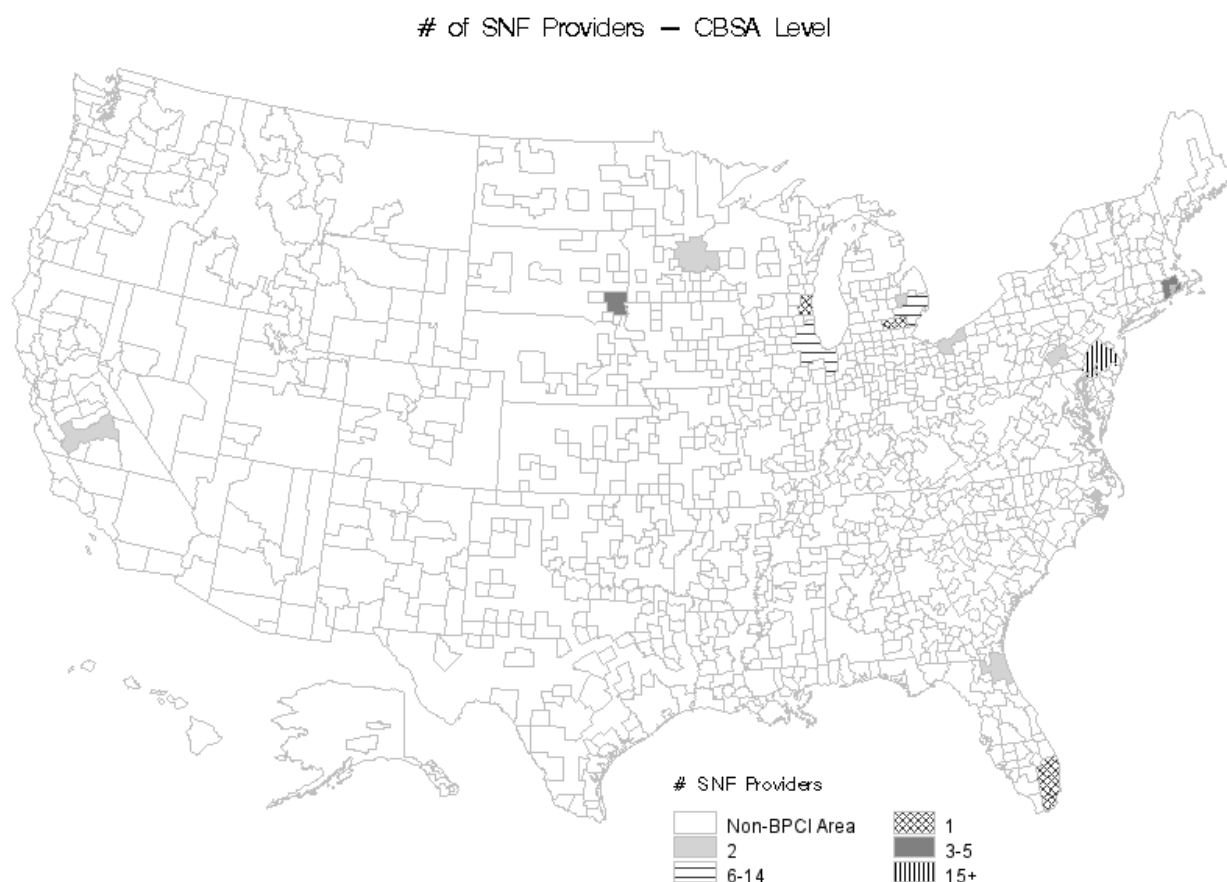
Source: Lewin analysis of CMS' BPCI database for all Q4 2013 – Q3 2014 BPCI participating SNF EIs.

---

[60] The market is defined as the Core Based Statistical Area (CBSA). Providers not located within a CBSA were assigned to the largest CBSA within their Hospital Referral Region (HRR). Non-BPCI markets represent all CBSAs that do not have a Model 3 BPCI participant. Areas of the country that are not in a CBSA are therefore not included in these non-BPCI markets.

Exhibit 91: Number of BPCI Participating IRFs by CBSA, Model 3, Q4 2013 – Q3 2014
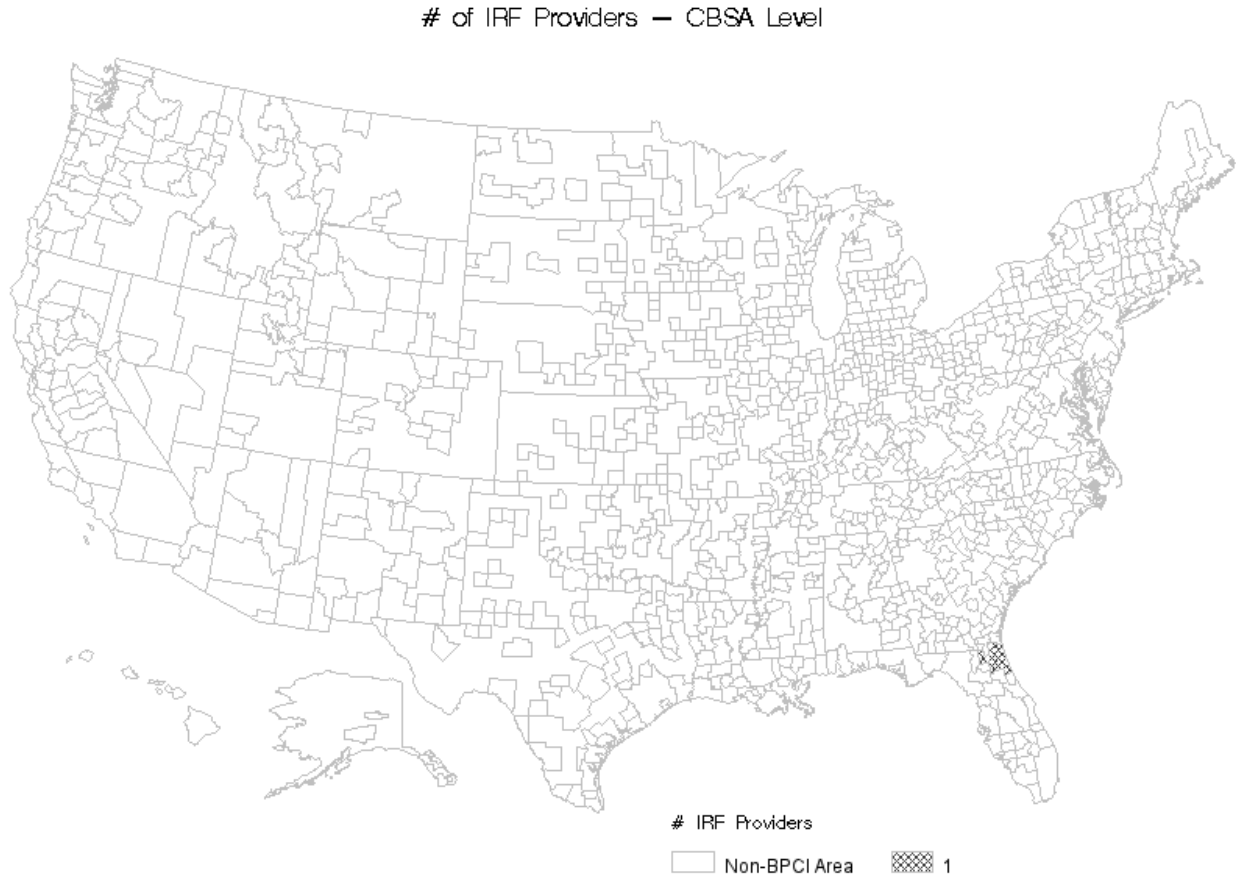
Source: Lewin analysis of CMS' BPCI database for all Q4 2013 – Q3 2014 BPCI participating IRF EIs.

Exhibit 92: Number of BPCI Participating HHAs by CBSA, Model 3, Q4 2013 – Q3 2014
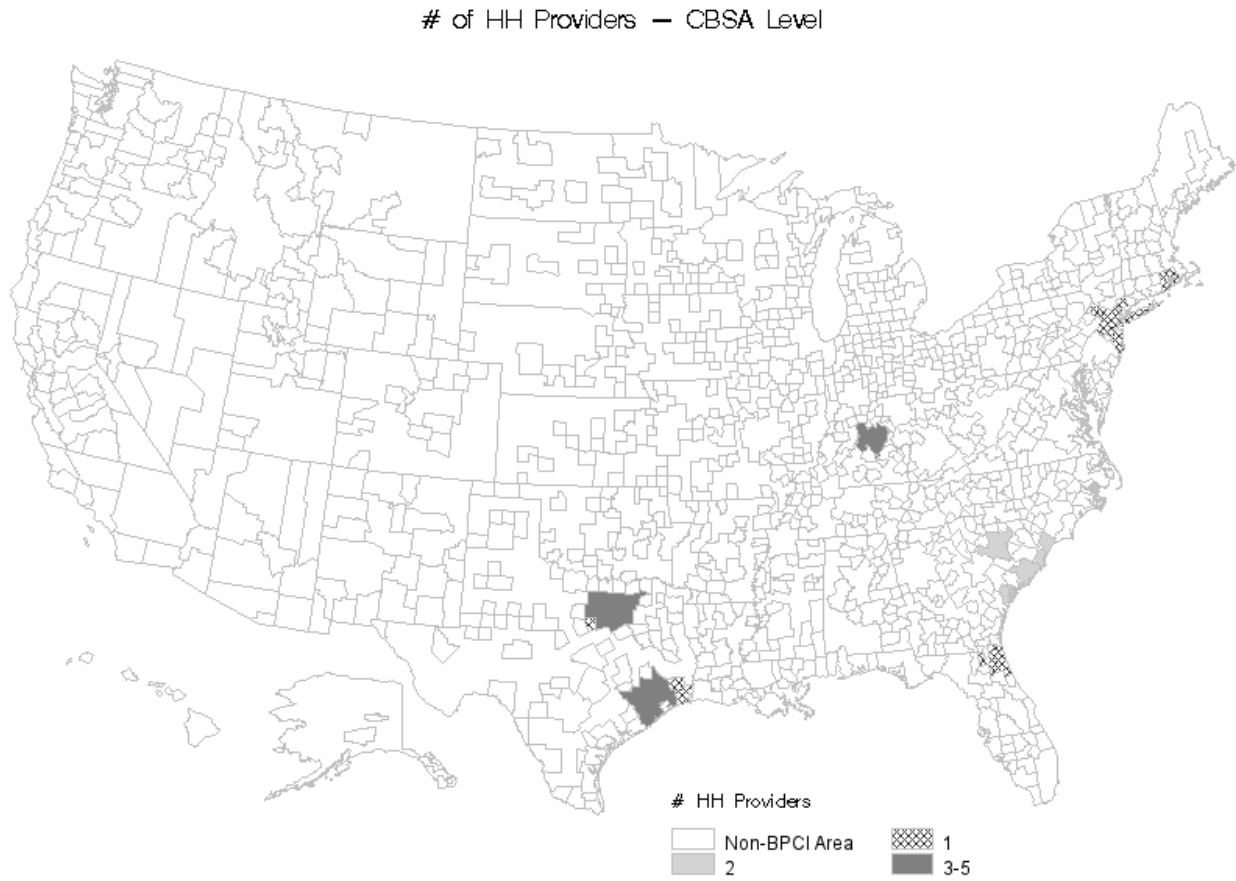
Source: Lewin analysis of CMS' BPCI database for all Q4 2013 – Q3 2014 BPCI participating HHA EIs.

Exhibit 93: Number of BPCI Participating LTCHs by CBSA, Model 3, Q4 2013 – Q3 2014

# of LTCH Providers — CBSA Level

LTCH Providers

Non-BPCI Area   1
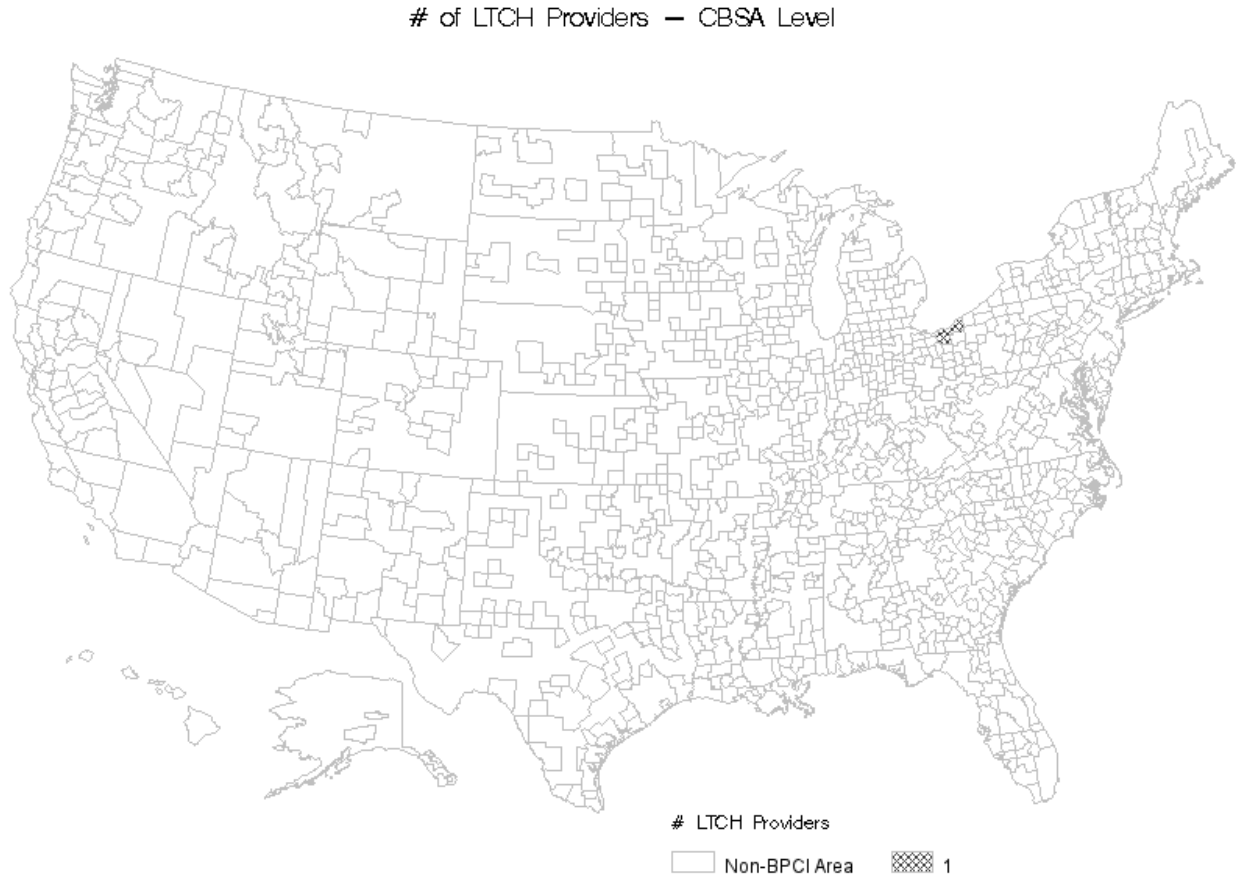
Source: Lewin analysis of CMS' BPCI database for all Q4 2013 – Q3 2014 BPCI participating LTCH EIs.

Exhibit 94: Number of BPCI Participating PGPs by CBSA, Q4 2013 – Q3 2014
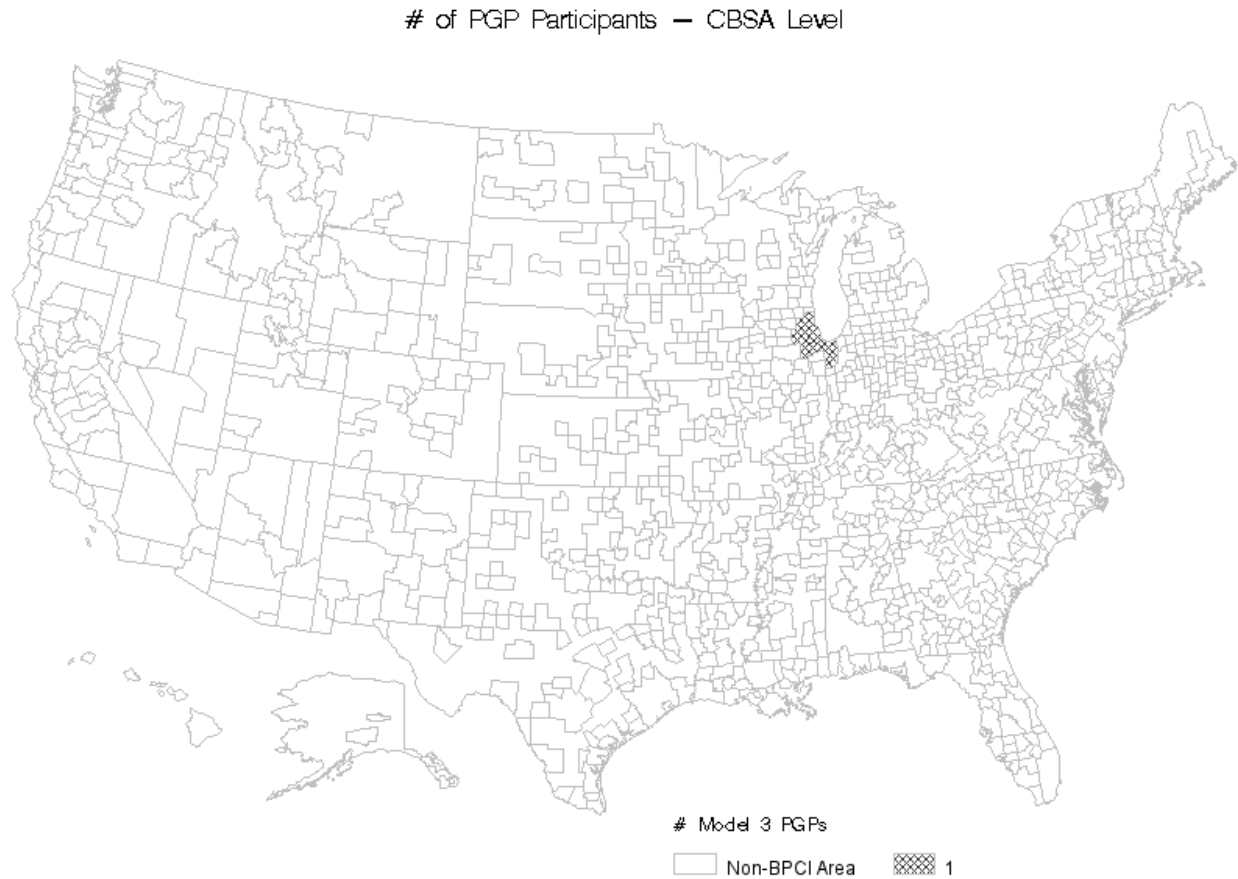
Source: Lewin analysis of CMS' BPCI database for all Q4 2013 – Q3 2014 BPCI participating PGP EIs.

Markets with BPCI-participating EIs differed from markets without BPCI EIs, as shown in Exhibit 95, which includes all Model 3 EIs. The average BPCI market penetration rate for SNFs in BPCI markets was 8.5%, meaning that BPCI-participating SNFs had 8.5% of the SNF beds in their markets. The average BPCI market penetration rate for HHAs was 18.7%. BPCI markets had high SNF competition (average Herfindahl index value of 0.07), which exceeded that of non-BPCI markets (0.32); this was also true of HHA competition in BPCI markets compared to non-BPCI markets (0.18 vs. 0.53). BPCI markets were more heavily populated (averaging about 2.8 million residents) compared with markets that had no BPCI-participating SNFs or HHAs (averaging about 238,000). On average, BPCI markets had a higher median household income than non-BPCI markets ($51,000 vs. $44,000), as well as more primary care physicians, specialists, and nurse practitioners per 10,000 residents.

Exhibit 95: Characteristics of BPCI Markets and Non-BPCI Markets, Model 3, Q4 2013-Q3 2014

Market Characteristics – Model 3	BPCI Markets N=26; 2.8% of Markets				Non-BPCI Markets N=916; 97.2% of Markets			
	Mean	Median	25th	75th	Mean	Median	25th	75th
BPCI Market Penetration – Hospital	7.1%	0.0%	0.0%	8.0%	2.3%	0.0%	0.0%	0.0%
Herfindahl Index - ACH	0.24	0.17	0.05	0.37	0.68	0.88	0.39	1.00
BPCI Market Penetration – SNF	8.5%	5.5%	0.0%	14.0%	0.0%	0.0%	0.0%	0.0%
Herfindahl Index - SNF	0.07	0.04	0.01	0.10	0.32	0.27	0.14	0.41
BPCI Market Penetration – HHA	18.7%	0.5%	0.0%	30.0%	0.0%	0.0%	0.0%	0.0%
Herfindahl Index - HHA	0.18	0.15	0.06	0.24	0.53	0.51	0.21	1.00
BPCI Market Penetration - IRF	3.8%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%
Herfindahl Index - IRF	0.40	0.32	0.00	1.00	0.12	0.00	0.00	0.00
Medicare Advantage Penetration	23.1%	22.4%	14.7%	26.5%	18.1%	15.2%	8.5%	24.2%
Population	2,822,329	1,118,877	347,962	4,285,832	238,846	71,953	39,396	162,368
Median Household Income	$50,787	$49,997	$44,989	$56,524	$44,021	$42,966	$38,422	$48,302
% Age 65+	14%	13%	12%	15%	15%	15%	13%	17%
PCP Per 10,000	7.7	7.8	6.1	8.9	6.4	6.2	4.7	7.6
Specialist Per 10,000	11.1	10.4	6.5	13.0	5.4	4.4	2.6	6.7
PA/NPs Per 10,000	7.1	6.6	4.9	8.2	6.1	5.6	3.9	7.6
SNF Beds Per 10,000	51.6	45.4	33.7	59.3	71.1	65.1	44.0	90.8
LTCH Beds Per 10,000	1.2	1.0	0.5	1.7	0.5	0.0	0.0	0.0
IP Rehab Bed Per 10,000	0.5	0.4	0.0	0.8	0.4	0.0	0.0	0.0
CAH Beds Per 10,000	0.2	0.0	0.0	0.1	1.8	0.0	0.0	1.3

Source: Lewin analysis of 2011 Medicare claims and 2011 Area Health Resource File (AHRF). Variable definitions are included in **Appendix N**.

3. Model incentive structure characteristics

a. Entry decisions

Twenty of the 94 Awardees participated in Model 3 during the first year of the initiative, representing 94 of the 227 EIs (1 PGP, 1 LTCH, 1 IRF, 28 HHAs, and 63 SNFs). The Awardee interviews and case studies provide insights into the reasons that organizations chose to participate in BPCI. We conducted 9 case studies and 18 Awardee interviews with Model 3 participants.

Nearly half of the respondents in the quarterly interviews said that the BPCI initiative provided an opportunity to learn about bundled payments and anticipated payment reform. This topic also arose during several of the Model 3 case studies conducted during the second year of BPCI. Interviewees at these Model 3 sites indicated that their interest in BPCI came primarily from the opportunities to learn, develop innovative approaches to care, and generate financial gains.

b.   *Episode and length selection*

Model 3 EIs participated in 46 out of 48 potential clinical episodes during the first year of BPCI. The average Model 3 EI participated in 19 clinical episodes. Non-surgical cardiovascular was the most frequently selected clinical group, with 95% of EIs participating in at least one clinical episode within this group. Roughly 95% of EIs participated in congestive heart failure, making it the most popular clinical episode among Model 3 participants. Chronic obstructive pulmonary disease and simple pneumonia were the next most common clinical episodes, chosen by 74% of EIs during the first year of the initiative. During the first four quarters, only eight clinical episodes had less than 10% participation from Model 3 EIs. Exhibit 96 summarizes the count of Model 3 EIs participating in each of the 48 clinical episodes by participant type.

Exhibit 96: Participation of Episode Initiators by Clinical Episode, Model 3, Q4 2013 - Q3 2014

Clinical Episode	Episode Initiators by Participant Type (N=94)					
	SNF (N=63)	HHA (N=28)	IRF (N=1)	LTCH (N=1)	PGP (N=1)	%
Non-surgical and Surgical: Gastrointestinal (GI)						
Esophagitis, gastroenteritis and other digestive disorders	38	0	0	0	0	40%
Gastrointestinal hemorrhage	38	0	0	0	0	40%
Gastrointestinal obstruction	38	0	0	0	0	40%
Major bowel procedure	38	0	0	0	0	40%
Total	**38**	**0**	**0**	**0**	**0**	**40%**
Non-surgical: Cardiovascular						
Acute myocardial infarction	39	15	0	0	0	57%
Atherosclerosis	37	0	0	0	0	39%
Cardiac arrhythmia	38	15	0	0	0	56%
Chest pain	38	15	0	0	0	56%
Congestive heart failure	61	27	0	1	0	95%
Medical peripheral vascular disorders	38	15	0	0	0	56%
Syncope & collapse	38	0	0	0	0	40%
Total	**61**	**27**	**0**	**1**	**0**	**95%**
Non-surgical Neurovascular						
Stroke	38	15	0	0	0	56%
Transient ischemia	38	0	0	0	0	40%
Total	**38**	**15**	**0**	**0**	**0**	**56%**
Non-surgical Orthopedic						
Fractures of the femur and hip or pelvis	45	1	1	0	0	50%
Medical non-infectious orthopedic	41	0	0	0	0	44%
Total	**45**	**1**	**1**	**0**	**0**	**50%**

Clinical Episode	Episode Initiators by Participant Type (N=94)					
	SNF (N=63)	HHA (N=28)	IRF (N=1)	LTCH (N=1)	PGP (N=1)	%
Non-surgical: Other Medical						
Cellulitis	43	0	0	0	0	46%
Diabetes	38	2	0	0	0	43%
Nutritional and metabolic disorders	38	0	0	0	0	40%
Red blood cell disorders	38	0	0	0	0	40%
Renal failure	38	0	0	0	0	40%
Sepsis	45	0	0	1	0	49%
Urinary tract infection	48	16	0	0	0	68%
Total	**50**	**18**	**0**	**1**	**0**	**73%**
Non-surgical: Respiratory						
Chronic obstructive pulmonary disease, bronchitis, asthma	50	19	0	1	0	74%
Other respiratory	45	15	0	1	0	65%
Simple pneumonia and respiratory infections	49	20	0	1	0	74%
Total	**50**	**22**	**0**	**1**	**0**	**78%**
Cardiovascular surgery						
AICD generator or lead	0	1	0	0	0	1%
Cardiac defibrillator	38	15	0	0	0	56%
Cardiac valve	38	15	0	0	0	56%
Coronary artery bypass graft	38	17	0	0	0	59%
Major cardiovascular procedure	0	15	0	0	0	16%
Other vascular surgery	38	17	0	0	0	59%
Pacemaker	43	0	0	0	0	46%
Pacemaker device replacement or revision	38	0	0	0	0	40%
Percutaneous coronary intervention	38	17	0	0	0	59%
Total	**43**	**20**	**0**	**0**	**0**	**67%**
Orthopedic surgery						
Amputation	5	1	0	0	0	6%
Double joint replacement of the lower extremity	3	1	1	0	0	5%
Hip & femur procedures except major joint	9	1	1	1	0	13%
Lower extremity and humerus procedure except hip, foot, femur	5	0	0	0	0	5%
Major joint replacement of the lower extremity	20	2	1	1	1	27%
Major joint replacement of the upper extremity	38	0	0	0	0	40%
Other knee procedures	43	0	0	0	0	46%
Removal of orthopedic devices	42	0	0	0	0	45%
Revision of the hip or knee	7	4	1	0	0	13%
Total	**58**	**5**	**1**	**1**	**1**	**70%**

Clinical Episode	Episode Initiators by Participant Type (N=94)					
	SNF (N=63)	HHA (N=28)	IRF (N=1)	LTCH (N=1)	PGP (N=1)	%
Spinal surgery						
Back & neck except spinal fusion	0	0	0	0	0	0%
Cervical spinal fusion	0	1	0	0	0	1%
Combined anterior posterior spinal fusion	0	0	0	0	0	0%
Complex non-cervical spinal fusion	38	0	0	0	0	40%
Spinal fusion (non-cervical)	5	1	0	0	0	6%
Total	**43**	**2**	**0**	**0**	**0**	**48%**

Source: Lewin Analysis of CMS' BPCI database, June 2015.
Note: The total number of EIs for a given clinical group will not add up to the total EIs participating in each of the clinical episodes within that group because EIs can participate in more than one clinical episode within the group.

In interviews, Model 3 participants shared their rationales for selecting episodes. Most participants noted that the decisions concerning episodes under BPCI were made by the organizations' administrative leadership. Interviewees also mentioned the importance of tapping into the expertise across their organizations (e.g., financial, physician, and nursing).

"Our experience with the CHF category, and we used gambling as an example, we felt after we looked at it that it was pretty much like rolling a dice to try to forecast whether we could be profitable or at great risk."

Some participants also selected episodes on the basis of patient volumes, targeting episodes that were perceived to be major drivers of cost. However, one site we visited in the past year had a different approach to episode selection. This EI specifically selected low-volume episodes because their convener informed them that their average cost in high-volume bundles was already very low, and therefore the site was unlikely to further improve upon those costs. As a result, the EI selected low-volume episodes for which their average costs were higher than state and regional averages, because they felt those presented the best opportunity to achieve savings through care redesign.

Other EIs offered alternative approaches to episode selection, such as selecting challenging episodes that offered the best opportunities for learning or selecting episodes that could be most effectively managed by a SNF. In one case, an Awardee that participated in the joint replacement episode added additional surgical episodes because they wanted all of their surgeons to contribute to their gains.

During interviews this past year, we started to hear from Awardees who terminated participation in BPCI episodes. This action was mentioned by two Model 3 respondents in regard to the congestive heart failure (CHF) episode. Respondents indicated that they terminated participation in this episode because it was too unpredictable and risky to manage successfully. They noted that it was difficult to accurately identify CHF patients, and in some cases the patients were not identified until after they had been discharged. Further, they stated that CHF patients are often chronically ill and have multiple comorbidities, making it difficult to prevent readmissions.

Through Q3 2014, Model 3 participants chose either 60- or 90-day episodes. Of the 94 EIs, 12 EIs (13%) chose 60-day episodes and 81 EIs (86%) selected 90-day episodes. Some of the Model 3 Awardees interviewed noted that the 90-day period was strongly recommended by CMS when

sites were applying to participate in BPCI. The recommendation from CMS was also cited by multiple EIs during case studies conducted in the past year. Another participant selected the 90-day episode at the recommendation of their AC, who conducted an actuarial analysis of the impacts of choosing 60-day or 90-day episode and showed that the latter would be preferable. Furthermore, interviewees reported that they had the impression that the application process for BPCI was competitive and they therefore wanted to put forward the strongest application possible; they believed that CMS preferred the 90-day episode length and therefore it was selected. Awardees also indicated that they selected 90 days because they felt that was the appropriate length of time to study their population and allow for their care redesign system to function as intended, a reason that was reiterated during the second year of the initiative.

During quarterly interviews and case studies, several participants expressed that they encountered challenges with respect to the way BPCI episodes were defined. Awardees and EIs indicated that episodes could include readmissions that were unrelated to the initial hospitalization or that the EI had no control over.

c. Conveners in BPCI

There are several roles participants may play in the BPCI initiative. The role determines whether the participant bears risk, initiates episodes, or serves an administrative function (e.g., as a non-episode initiating convener). There are three different convener roles in BPCI: FC, DAC, and AC. FCs are non-episode initiating and non-risk bearing participants. These participants typically serve administrative and technical assistance functions on behalf of DACs and DAs. ACs and DACs serve similar functions, though DACs participate in BPCI under a FC. ACs and DACs may also initiate episodes, and they assume the financial risk on behalf of their EIs.

> *"This was really the first opportunity we saw where post-acute care could make a difference in how the reimbursement is done with Medicare. And so, that's part of the reason we jumped onto it. We thought it was an opportunity to be a leader in how this care redesign was going to occur with new reimbursement systems, and we just felt we were up to the challenge of doing that."*

From Q4 2013 through Q3 2014, there were four FCs and eight ACs participating in Model 3. All EIs that participated in Model 3 during the first year of BPCI were within a convener structure. In addition, over two-thirds of the EIs participated under three Awardee Conveners; these ACs accounted for 74% of the episodes initiated during the first year of the initiative.

To better understand the role of these conveners in BPCI, one set of quarterly interviews conducted this past year focused on FCs and their reasons for joining the BPCI initiative in that role. Representatives from all four Model 3 FCs participated in these interviews.

Those interviewed presented various reasons for joining BPCI as an FC. One respondent noted that having access to the data from CMS presented a "huge opportunity" to educate their orthopedic surgeons about their patients' outcomes. For another FC, BPCI was seen both as a chance to enhance their reputation with local hospitals and as an opportunity to save money on readmissions. One respondent viewed BPCI as an opportunity to be a leader in care redesign, and they believed that joining as a FC would allow their EIs to receive the benefits of redesigning care in their facilities.

Similar to Model 2 FCs, all Model 3 FCs interviewed believed that their primary role in BPCI was to educate their participating EIs in payment reform. Interviewees also described their role as a consultant and coordinator. In terms of care redesign, one FC stated that it facilitated a collaborative process among its members. The FC has arranged opportunities for participants to discuss specific tools and methods to improve care quality and patient outcomes. Another FC described a top-down approach, which included the development and standardization of efforts to reduce readmissions. Some FCs also provided guidance on episode selection. Those interviewed described the process of analyzing existing data and sharing this with their participants, though the final episode selection decision has generally remained with the participants.

According to Model 3 participants, ACs and DACs play less of a consultant or facilitator role, focusing more on designing, developing, and implementing care protocols. One AC hired new staff and built the staffing model for its EIs. This AC also spent significant time onsite at its EIs' facilities, interacting with beneficiaries and EI staff. According to one of their EIs, this AC made a concerted effort to find the correct balance between being overly involved and too passive. Further, this AC is very active in patient tracking and has a process to arrange beds, transportation, and other beneficiary needs when transitioning from one facility to another. According to its EIs, this AC was not well-known among the Model 3 participants and has worked to integrate its staff into each EI's operations.

Other Model 3 ACs and DACs focus less on the implementation of care redesign, opting instead to provide administrative support and guide episode selection. According to one Model 3 participant, its AC is an outside entity unrelated to its EIs, but played an integral role in selecting the EI's bundles. Further, the AC provided data analysis and administered the BPCI program overall. Further, this AC ensured financial reconciliations were accurate and shared those results with their EIs. The EI did express that it wanted its AC to be more involved in care redesign, such as providing assistance with data analysis and in identifying BPCI patients.

d. Partners

In interviews, Model 3 participants described seeking partnerships with a variety of entities. Model 3 participants continue to partner with other PAC facilities. Some participants described owning PAC facilities, with which they had developed strong partner relationships prior to joining BPCI. The goal of these partnerships with PAC facilities is to ensure downstream quality (i.e., work with rehabilitation partners to improve a patient's functional status), track patients after discharge to reduce readmission rates, and to stratify patient risk. Model 3 participants have also partnered with hospitals from which they receive their BPCI patients. One EI described having monthly meetings with their partner hospital to review cases that were at risk for readmission or that had already been readmitted. Further, this EI described the importance of this relationship in terms of data sharing and facilitating access to EMRs that could better identify BPCI patients.

"I think that [the hospitals, nursing homes, and home care entities] are all talking the same language and communicating that same information across the care continuum… The navigators are reinforcing that and collaborating with the primary care physicians to hopefully decrease the readmission rate."

Model 3 participants also continue to partner with third-party administrators and contractors to assist with data management, BPCI program administration and oversight, and gainsharing payment calculation and distribution. Some interviewees have formal partnerships with consulting firms that

analyze data and assist with patient identification and medication reconciliation. During the site visits, some EIs described how their convener plays this third-party administrator role. One EI explained that their AC analyzes EMR data to assist with patient identification and to track patients post-discharge.

Model 3 participants described several benefits as a result of these partnerships. One Awardee noticed a more collaborative relationship with their partner organizations. This Awardee described a workgroup with staff from five BPCI-participating facilities that has since improved communication and overall patient care. Another Awardee described the benefits of having internal partners, such as a hospital or other downstream PAC providers within the same network. These relationships allowed for organizational support and common care strategies and protocols. One Model 3 Awardee explained that sharing care protocols resulted in a cost benefit to Medicare, improved care coordination, and could ultimately reduce readmissions.

"We are asking the orthopedic surgeon to give us more detailed information about the care he feels would benefit that patient in the SNF. And, realistically an orthopedic surgeon will not take the time to do that."

Model 3 participants still noted challenges associated with partnerships. Participants continued to encounter difficulties working with physicians from different provider organizations and find that referral sources, especially those out of market, are often unwilling to collaborate. When speaking to their partnerships with referring hospitals, Model 3 participants noted that obtaining data from hospitals to facilitate patient identification was often a challenge. Participants described hospitals as being unresponsive to requests for information and some cited legal concerns related to HIPAA. Others noted challenges with interoperating with or using information contained in hospitals' EHRs, or inconsistency between a hospital's assigned DRG and the PAC facility's DRG. One Model 3 participant cited poor engagement among external partners such as physicians and hospitals as their greatest challenge.

In some markets, there were also challenges related to the increased number of partner hospitals that joined BPCI as Model 2 Awardees. One such challenge is identifying which facility qualifies as the EI for a given patient. Describing this situation, one Awardee said, "It has become increasingly confusing as to who 'owns' a patient's costs and/or benefits from the savings recouped." Model 3 Awardees noted another challenge related to when a partner hospital joins BPCI; in order to cut down on the cost of their own episodes, Model 2 hospitals tend to reduce SNF utilization and send their patients to home health, thus affecting the Model 3 Awardees' patient volume.

e. Waiver use

To use the beneficiary incentives or gainsharing waivers, an Awardee must describe its plans for waiver use in its IP. Other waivers are available to all Model 3 participants. Based on a review of Awardee IPs, 13 Awardees active during Q3 2014 were allowed to use one or both of these

waivers. Eleven EIs do not intend to use either waiver.[61] Exhibit 97 provides the count of Model 3 EIs who used each of the four waivers available to Model 3 participants.

Exhibit 97: Participation of Episode Initiators in BPCI Waivers, Model 3, Q3 2014

Model 3 Waivers	Model 3 EIs (N=94)			
	Model 3 EIs allowed to Use Waiver		Model 3 EIs that Used Waiver	
	N	%	N	%
Beneficiary Incentives	36	38.3%	17	18.1%
Gainsharing	78	83.0%	NA*	NA*
Telehealth	NA**	NA**	0	0.0%
Home Visit	NA**	NA**	0	0.0%

*No data are available regarding use of the gainsharing waiver.

** The home visit and telehealth waivers are available to all Model 3 Awardees without specifying it in their Implementation Protocols.

Sources: Lewin analysis of Awardee Implementation Protocols for Q3 2014 BPCI participants, Medicare claims data for episodes initiated Q4 2013 – Q3 2014, and Awardee-submitted data Q4 2013 – Q3 2014.

Note: The 94 EIs in Q3 2014 are distributed among 20 Model 3Awardees. EIs include SNF, HHA, PGP, IRF, and LTCH. Designation of the use of the telehealth and home visit waivers in the Implementation Protocol is not required. Due to a delay in collecting information from Awardees regarding gainsharing, we are unable to determine how often this waiver has been used during the first year of the initiative.

Beneficiary incentives

The beneficiary incentive waiver allows the EI to offer a service or product to a beneficiary that is related to the episode, but that is not typically covered by Medicare. As displayed in Exhibit 97 above, in Q3 2014, 36 EIs (roughly 38%) signed up for the beneficiary incentives waiver, allowing EIs the opportunity to provide these incentives to BPCI beneficiaries. We analyzed Awardee-submitted data that provided details about the beneficiary incentives that were distributed to beneficiaries between Q4 2013 and Q3 2014. Twenty-four Model 3 EIs reported that they distributed incentives to 186 BPCI beneficiaries during this time period. These incentives ranged in value from $13.29 to $1173.20 during Q4 2013 through Q3 2014. Exhibit 98 provides a description of the distribution of beneficiary incentives by category for incentives that were provided in Q3 2014. In Q3 2014, as in the first four quarters of the initiative overall, incentives categorized as equipment were most common among Model 3 EIs that provided beneficiary incentives.

---

[61] For a description of the waivers, see section I.A.2 BPCI Initiative.

Exhibit 98: Beneficiary Incentives Waivers Distributed by EIs to beneficiaries, Model 3, Q3 2014

Incentive Description	Awardees allowed to use incentives	EIs allowed to use incentives	Awardees that provided one or more incentive	Episodes receiving one or more incentive	Average cost per incentive ($)
Transportation	1	1	1	6	$275.12
Equipment	3	7	3	37	$43.56
Home care/home visits	1	1	1	13	$347.58
Living arrangement services	1	1	1	2	$94.50
Telehealth/technology	7	7	2	7	$96.57
Wellness program/resources	2	18	0	0	NA
Medication management tools	1	1	0	0	NA

Source: Lewin Program Adherence Report based on analysis of Awardee-submitted data regarding disseminated beneficiary incentives in Q3 2014.

The use of beneficiary incentives was also discussed during our site visits. Specifically, some of the EIs described their concerns with the ethics and legality of the beneficiary incentive. One EI, that is not offering any beneficiary incentives, was specifically concerned about the legality of offering transportation to use the services they provide. This EI hoped to learn from other participants with experience in beneficiary incentives related to transportation. Other EIs offered electronic tablets and personal emergency response systems to maintain contact with the patient and provide real-time assistance in the event of a complication. During our site visits, one EI noted that they did not offer beneficiary incentives. This EI noted that because these products and services were not offered to all patients at no cost, there were potential ethical conflicts of providing this benefit for one group of patients and not others.

Gainsharing

Gainsharing enables participants to share any savings, with limitations, generated under BPCI with its gainsharing partners, including its EIs. According to the Awardee IPs, in Q3 2014, 78 EIs (83%) have signed up for gainsharing (Exhibit 97). Out of those EIs who are participating in gainsharing, 12 intend to gainshare with physicians.[62]

Awardees continue to view gainsharing as a tool to incentivize cooperation among upstream and downstream partners. During a quarterly interview, one Awardee emphasized the accountability to other providers that is implicit in gainsharing arrangements.

> *"[Gainsharing] is giving them incentive to hold one another accountable to the performance and quality changes."*

The Awardee believes gainsharing will encourage compliance with quality metrics and enhance patient monitoring after discharge. On several site visits, participants noted that they opted for simple and manageable gainsharing methodologies. Some conveners plan to share gains with the EIs directly, rather than involving other partners in the continuum of care. Of those Awardees that

---

[62] Due to a delay in collecting information from Awardees regarding gainsharing, we cannot comment on how often this waiver has been used during the first year of the initiative.

gainshare with other facilities, many conducted thorough reviews of the providers in their market to narrow their pool of prospective partners. To focus their search, one Awardee developed a list of requirements for their gainsharing partners, including an acceptable rate of readmission.

Of those participants that chose not to gainshare, many attributed the decision to the complexities of their market. Several EIs noted that their market is saturated with a particular type of facility, giving potential gainsharing partners little incentive to coordinate care. One BPCI-participating HHA mentioned that neighboring hospitals were disinterested in partnering with their relatively small facility because it receives only a fraction of the hospital's patients. Awardees were also deterred by the financial cost and time required to execute gainsharing agreements. One EI anticipated that gains would not compensate for the costs of implementation. Further, this EI noted that the time lapse between an implemented care redesign intervention and corresponding gainsharing payment would make it difficult to sustain positive changes in care provision.

The convener structure and relationship also impacts the BPCI participant's decision to gainshare. This factor was especially prevalent among Model 3 EIs that participated in case studies, all of which are members of a broader FC or AC group. The majority of these sites noted that the gainsharing agreement was not executed at the EI level. Gainsharing EIs share in gains but bear no downside risk; some EIs, however, are charged fees by their Conveners to cover the costs of data analytics and program administration.

Telehealth and home visit waivers

Telehealth services usually can be covered by Medicare where the originating site is one of eight health care settings that is located in a geographic area that is designated as a rural health professional shortage area, or in a county that is not included in a Metropolitan Statistical Area. The geographic requirement is waived under the telehealth waiver, thereby making telehealth services available to beneficiaries in urban and other areas that would not usually qualify.

The post-discharge home visit waiver waives the usual requirement for direct supervision for such visits and allows Awardees to provide home visit services after discharge from an EI to beneficiaries who are not otherwise eligible for home health services. Home visits must be furnished by licensed clinical staff under general supervision, billed by the supervising practitioner, and provided not more than once in a 30-day episode, twice in a 60-day episode, or three times in a 90-day episode.

Because the home visit and telehealth waivers are available to all Model 3 Awardees without requesting it through the IPs, we do not have any information about participants' intentions to use these waivers. However, an analysis of Medicare claims data suggests neither waiver was used for any BPCI episode in the first four quarters of the initiative.

4. Care redesign and cost saving strategy characteristics

During the first data submission period, Model 3 EIs reported if they participated in five types of care redesign activities and the status of each intervention during 2014. As shown in Exhibit 99, the majority of SNF EIs reported participating in all five types of care redesign; only one SNF reported not participating in interventions related to the redesign of care pathways. The majority of SNF EIs that reported participating in a care redesign activity indicated that their activity was fully operational (67-81%).

Exhibit 99: Participation and Status of Care Redesign Interventions, Model 3 SNFs, 2014

	Redesign of Care Pathways	Enhancement in Care Delivery	Patient Activation, Engagement, and Risk Management	Care Coordination	System Changes to Support Care
Percentage of EIs that indicated participation	98%	100%	100%	100%	100%
Care redesign status (among participating EIs)					
In planning stage	8%	3%	11%	3%	3%
Implementation started	25%	26%	15%	21%	16%
Fully operational	67%	71%	74%	76%	81%
Completed & terminated	0%	0%	0%	0%	0%

Source: Lewin analysis of Awardee-submitted data collected February 2015 for 2014 Model 3 BPCI participants.

Similarly, all 28 Model 3 HHA EIs reported participating in the five types of care redesign activities during 2014, as shown in Exhibit 100. The majority of EIs participating in a care redesign activity reported that their care redesign activity was completed and terminated (54%).

Exhibit 100: Participation and Status of Care Redesign Interventions, Model 3 HHAs, 2014

	Redesign of Care Pathways	Enhancement in Care Delivery	Patient Activation, Engagement, and Risk Management	Care Coordination	System Changes to Support Care
Percentage of EIs that indicated participation	100%	100%	100%	100%	100%
Care redesign status (among participating EIs)					
In planning stage	0%	0%	0%	0%	0%
Implementation started	4%	39%	39%	4%	7%
Fully operational	43%	7%	7%	43%	39%
Completed & terminated	54%	54%	54%	54%	54%

Source: Lewin analysis of Awardee-submitted data collected February 2015 for 2014 Model 3 BPCI participants.

a. Care redesign and care coordination

The case studies and quarterly interviews provided more details on how participants were implementing care redesign activities across the five categories described above. Many participants described activities related to risk assessment, care coordination, and patient education.

The majority of Model 3 Awardees that were interviewed over the past year have introduced risk assessment as part of their care redesign strategy. For example, one large convener introduced a three-level system that assigns an acuity level for every BPCI patient (e.g., a "red" patient has a high likelihood for readmission; a "yellow" patient has a history of acuity or readmissions, but not as recently as a "red" patient). While this convener did not use care protocols specific to the different risk statuses, a higher risk level would trigger different "assignments" for the case management staff.

Nearly all Model 3 participants that we spoke with during quarterly interviews and case studies indicated that care navigators/coordinators/case managers were a critical component of their

care redesign efforts. These positions, frequently filled by nurses, were often created specifically for BPCI. The roles of these staff varied by participant and included helping patients schedule appointments, filling prescriptions, coordinating home support, tracking patients after discharge, and communicating with primary care physicians. One large Model 3 convener hired more than 50 new staff to serve as case managers at its EIs. For some participants, navigators were used for all BPCI patients; for others, case management was only used for patients who were deemed high-risk. One Model 3 Awardee noted that they were increasingly seeing patients return at around the 100 day mark, following the end of the bundle period and case management services, which shows the importance of case management for Medicare patients.

Model 3 participants also emphasized the importance of improved patient education. Respondents from the majority of Model 3 case studies confirmed that education early in a patient's stay was enhanced – "focusing on the front-end with education rather than at discharge." Some of the goals of this education included: informing both patients and their family members how to take care of themselves (e.g., coaching them on what to ask their PCP), setting

> *"We're providing education from the day the patient arrives, starting with discharge planning as they are getting off the stretcher. Setting an expectation with the patient that their estimated length of stay will for a simple knee joint would be 7 to 10 or 10 to 14 days, as opposed to their understanding that they have 100 days of the Medicare skilled [nursing facility] benefit."*

expectations for their expected length of stay, and helping them to change behaviors that could lead to an ER visit. One large convener noted that their staff conducts some education through home visits. Model 3 participants indicated that these efforts were important to achieving their goal of reduced readmissions.

Despite this enhanced education, patients' expectations for post-acute care presented a challenge for Model 3 participants. Several noted that patients often feel entitled to 21 SNF days and expect high-end facilities with full amenities. Shifting this mindset has been a challenge, particularly for PGPs who see reducing SNF LOS as a method of cutting costs. One PGP described the challenge in the following way: "There is an abundance of SNFs offering spa-like services to a clientele that expects it (such as massage and hotel-like amenities) and more high-end SNFs are being built. Fifty percent of elective patients have already 'booked a room' before their surgery – so we have no opportunity to divert them from SNF care." Other participants described patients who are resistant to changing their perspectives of PAC and perceive hospitals or the emergency department as the only places to receive acute care.

Model 3 participants also faced internal challenges related to patient discharge procedures, reporting that a large cultural shift among their staff was necessary for bundled payments to be successfully implemented. Awardees highlighted that revising their approach to discharge planning necessitated clear communication from the first day a patient was admitted. One LTCH described the transformation required to change their traditional approach of sending patients to a SNF prior to being discharged home; as a result of BPCI, their staff now challenge that assumption, asking "why can't they go home?"

Another challenge reported by Model 3 participants was identifying which patients will be BPCI beneficiaries. Awardees noted that it was difficult to know who qualified as a BPCI patient because the inpatient DRG, which determines whether a patient is included in BPCI, was not always known at the time of admission to PAC. Confirmation of the anchor DRG often took

months, which forced Awardees to deliver care redesign based on their best guess of who would later be identified as a BPCI beneficiary. Oftentimes, inaccurate identification led to a misallocation of BPCI resources.

b. Cost savings strategies

The most common cost-saving strategies used by Model 3 participants focused around two areas: reducing PAC utilization and reducing readmissions. Reducing PAC utilization as a means to achieve savings has been a consistent theme since the beginning of the initiative.

> *"Prior to the BPCI program, we assumed that patients would go to SNF or acute rehab. Now we're really looking at getting people home—more opportunities for that."*

According to Awardee interviews, Model 3 Awardees believe that reducing PAC utilization was a key factor for success in the BPCI initiative. Nearly every Model 3 Awardee that was interviewed during case studies and quarterly calls mentioned that better managing

PAC utilization was part of their plan to reduce costs. Awardees had different strategies for achieving this. Some Awardees indicated that their goal was to divert patients to a HHA and reduce the number of days of SNF care. One Awardee, a LTCH, began to identify patients at admission who were appropriate for

> *"As time went on and we evolved as a program, we realized our focus is mostly to prevent readmission once patients are in the home and how to communicate with the family to improve this outcome."*

discharge to home. As a result of BPCI, their therapists provided early and continual ambulation to prepare patients for a successful discharge and made early contact with family members.

> *"Our program is really about preventing avoidable readmissions. It's not a length-of-stay play. We're really trying to make sure that these patients are in a good place when they're discharged from the skilled nursing facility and that they have the right tools to be successful in the community."*

Other Model 3 participants are more focused on reducing readmissions. In site visits and quarterly calls, we heard of different approaches to achieving this goal. Some participants started additional training and education for their nurses, making them more aware of the most common causes of readmissions and the symptoms of those causes. Others began doing a root cause analysis for all readmissions to identify areas for

improvement, such as enhancing clinical pathways. In some cases, the goal of reducing readmissions was at odds with the goal of reducing PAC utilization. Several participants stated that length of stay was not their primary concern; rather, they preferred patients to have additional days in a SNF if it reduced the risk of readmission. Generally speaking, readmission management appeared to be an area of success, as all of the Model 3 sites that were visited in the past year reported a decrease in their readmission rates since joining the initiative.

Model 3 participants also reported challenges with cost savings. Several noted a shift in the type of patients they were admitting, with several participants reporting an unexpected increase in the number of complex patients (e.g., sepsis, pneumonia, and those with multiple comorbidities). This introduced challenges related to bundling care. One such challenge was the drain on resources, including staff time, greater length of stay, and higher expenses for treatment. During quarterly interviews, one Awardee described the conflicts that they experienced, saying, "keeping a patient here longer … is the right thing for the patient but you lose money – you are stuck." Medically complex patients required more intensive therapy and had an increased risk for re-hospitalization, according to interviewees. These challenges involved in caring for patients with

chronic diseases made it difficult for participants to reduce costs for these episodes while maintaining the quality of care.

During the past year we completed interviews with participants that withdrew from the program as of June 2015; these included a single Model 3 EI. The concerns regarding costs summarized above were similar to what was heard in the interview. The participant indicated that the main reason they exited the initiative was that they faced operating costs related to BPCI that were greater than they could afford (approximately $150,000 per quarter). The interviewee also indicated that they faced challenges in identifying patients' DRGs, which further led to their decision to withdraw from the initiative.

B.   Impact of BPCI

This section presents the first year impact estimates of Model 3 BPCI (SNF and HHA) episodes on payments, utilization, and quality of care based on episodes that were initiated during the first four quarters of the BPCI initiative (Q4 2013 through Q3 2014).[63] For Model 3, we performed the quantitative analysis separately for each type of EI (SNF and HHA) because the patients and, as a result, the patterns of care, differ across the settings. [64]

We present results for the key outcomes across all Model 3 SNF episodes and all Model 3 HHA episodes in this section, and then separately by clinical episode group in Sections IV.D through IV.H. The exhibits present the estimated differential change in outcomes for patients receiving care from BPCI providers, between the baseline and the intervention period, relative to the same change for the patients receiving care from providers in a comparison group (DiD). See Section II.D.2 for additional details on the statistical approach.

1.   *Change in Medicare standardized allowed payment amounts*

a.   *Model 3 SNF-initiated episodes*

This section summarizes the results for the key payment outcomes for Model 3 SNF EIs relative to the comparison group. We calculated the Medicare standardized allowed payment amount by type of service as well as across all services for two measurement periods. The first measurement period is the length of the episode (30-, 60-, or 90-days). We present the total payments included in the bundle, and the total payments not included in the bundle, by bundle length. We also show results for total payments outcomes for the 60 days after the end of the episode (post-bundle period) and for the 30 days prior to the start of the episode (pre-bundle period). The second measurement period is the qualifying hospitalization and the 90 days after the hospital discharge (90 day PDP), regardless of the episode length. See **Appendix L** for detailed outcome definitions.

---

[63] Please note that pre-bundle and post-bundle period payment outcomes and patient-assessment based outcomes are reported with a one-quarter delay and therefore are based on episodes that were initiated during the first three quarters of the BPCI initiative (Q4 2013 through Q2 2014).

[64] There was only one IRF and one LTCH participating in BPCI during the first year of the initiative. Given the small number of patient episodes for a single BPCI participant for each EI type, we do not present the results for the IRF and LTCH participants.

Exhibit 101, presents the direction and point estimate of the DiD estimate for the key payment outcomes and each clinical episode group for Model 3. It should be noted that the columns are not mutually exclusive. The surgical sample includes all surgical episodes, including the orthopedic surgery clinical episode group. Similarly, the nonsurgical sample includes all nonsurgical episodes, including nonsurgical respiratory and non-surgical cardiovascular. Therefore, the results of the larger surgical and nonsurgical groups are largely driven by the three clinical episode groups displayed in the table. The top panel shows payment outcomes that were measured based on episodes initiated during the first year of the initiative. The bottom panel shows payment outcomes that were measured based on episodes initiated during the first three quarters of the initiative. These include post-bundle and pre-bundle payment outcomes, which were measured based on episodes initiated during the first three quarters of the initiative in order to allow for sufficient claims run out. Orange highlighted cells represent a statistically significant decrease relative to the comparison group. Green highlighted cells represent a statistically significant increase relative to the comparison group. The baseline and intervention estimates, for both BPCI and the comparison group, can be located for key outcomes in Sections IV.D through IV.I.

Model 3 SNF EIs were successful in reducing standardized payments for SNF services, but the lower payments were offset by payments for other services by the end of the 90 day post-discharge period. Total (Part A and B) standardized allowed payment amounts for Model 3 BPCI SNF EIs was similar for BPCI beneficiaries and beneficiaries treated by comparison providers. The total standardized allowed payment included in the bundle definition declined across all other clinical episode groups, however, these differences were not statistically significant relative to the comparison group. However, we found declines in SNF payments. For orthopedic surgery episodes, the decline in the SNF standardized allowed amount during the, 90-day PDP was statistically significant relative to the change for the comparison group (-$2,101). The change for non-surgical cardiovascular episodes in SNF standardized allowed payment during the 90-day PDP was also statistically significant relative to the comparison group (-$2,708). Because orthopedic surgery and non-surgical cardiovascular episodes dominate the surgical and non-surgical clinical episode groups, respectively, the two aggregate clinical episode groups also had statistically significant declines in the SNF standardized allowed payment, 90-days PDP relative to the comparison groups (-$1,733 and -$1,330, respectively). Finally, during the first three quarters of the initiative, neither the change in the total amount included in the pre-bundle period, nor the change in the total amount in the post-bundle period, were statistically significant relative to the comparison group. Therefore, there is no reason to believe that providers were shifting costs outside of the bundle period to reduce costs under BPCI.

Exhibit 101: Diff-in-Diff Estimate for Allowed Payment Outcomes, by Clinical Episode Group, Model 3 SNFs, Baseline to Intervention

Measure	Length	Surgical	Nonsurgical	Orthopedic Surgery	Nonsurgical Respiratory	Nonsurgical Cardiovascular
Number of episodes initiated Q4 2013 – Q3 2014		1,879	5,467	1,303	988	1,087
BPCI Cumulative to Date (Q4 2013 – Q3 2014)						
Standardized allowed amount (Part A & B), IP through 90-day post-discharge (PD) period		-$568	-$1,069	-$1,553	-$854	-$969
Standardized allowed amount in bundle definition	60	-$1058		-$1,058		
Standardized allowed amount in bundle definition	90	-$884	-$851	-$1,285	-$920	-$848
Standardized allowed not included in bundle definition	90	$53	$40		$41	-$265
SNF Standardized allowed amount, 90-day PD		-$1,733	-$1,330	-$2,101	-$806	-$2,708
First Three Quarters of BPCI (Q4 2013 – Q2 2014)						
Standardized allowed amount Part A & B, Days 1-30 Post bundle payment (PBP)	60	-$126		-$130		
Standardized allowed amount Part A & B, Days 1-30 PBP	90	-$369	-$269	$370	-$18	-$559
Standardized allowed amount Part A & B, Days 31-60 PBP	60	-$450		-$450		
Standardized allowed amount Part A & B, Days 31-60 PBP	90	-$213	-$291	-$37	-$329	-$398
Standardized allowed amount Part A & B, pre-bundle period	90	$767	-$226		-$726	-$133

Note: Statistical significance at the 0.05 level is indicated by orange and green shaded cells. Orange indicates the DiD estimate was negative and statistically significant; green indicates the DiD estimate was positive and statistically significant. The results presented in this table represent episodes initiated in the first four quarters of the BPCI Initiative (Q4 2013 through Q3 2014). Please note that pre-bundle and post-bundle period (PBP) payment outcomes are reported with a one-quarter delay. Many of the payment outcomes are stratified by length of episode. A blank cell indicates that the outcome cannot be presented, either due to insufficient sample size or the type of episodes initiated during the time period. It should be noted that the columns are not mutually exclusive. The surgical sample includes all surgical episodes, including the orthopedic surgery clinical episode group. Similarly, the nonsurgical sample includes all nonsurgical episodes, including nonsurgical respiratory and nonsurgical cardiovascular. Therefore, the results of the larger surgical and nonsurgical groups are largely driven by the three clinical episode groups displayed in the table.

Source: Lewin analysis of Medicare claims and enrollment data for episodes that began Q4 2011 through Q3 2014 for BPCI providers and a comparison group.

We calculated a number of utilization outcomes to understand the changes that contributed to the payment outcomes summarized above. Key utilization measures include average inpatient length of stay (LOS), number of HH visits post qualifying hospitalization discharge, number of days during the 90 days post qualifying hospitalization discharge in each institutional setting (SNF, IRF, and ACH for readmissions)[65], and total number of days in any institutional setting after the qualifying hospitalization discharge. Number of days of PAC use and number of HH visits are for patients who had at least one day in the respective settings during the 90-day post-discharge period. See **Appendix L** for detailed outcome definitions.   The baseline and intervention estimates, for both BPCI and the comparison group, can be located for key outcomes in Sections IV.D through IV.I.

For all clinical episode groups, except non-surgical respiratory, there was a statistically significant decline in the number of SNF days relative to the comparison group, which is consistent with the payment measures described above (Exhibit 102).   For non-surgical cardiovascular clinical episodes, there was a statistically significant decline in the SNF days for BPCI episodes relative to comparison episodes (-4.5), and in the number of institutional days (-3.3). For the orthopedic surgery clinical episode group, the number of SNF days and the number of institutional days also had a statistically significant decline relative to the comparison group (-3.4 and -3.5 respectively). In the orthopedic surgery clinical episode group, this decline was accompanied by a statistically significant increase in the number of home health visits relative to the comparison group (2.0 visits).

---

[65] There was insufficient sample size to examine average LTCH length of stay.

Exhibit 102: Diff-in-Diff Estimate for Utilization Outcomes, by Clinical Episode Group, Model 3 SNFs, Baseline to Intervention

Measure	Surgical	Nonsurgical	Orthopedic Surgery	Nonsurgical Respiratory	Nonsurgical Cardiovascular
Number of episodes initiated Q4 2013 − Q3 2014	1,879	5,467	1,303	988	1,087
Acute Inpatient Care LOS	0.1	-0.1	0.0	-0.3	0.1
Number of HH Visits, 90-day PDP[1]	2.0	-0.3	2.2	0.4	0.6
Number of Institutional Days, 90-day PDP[1]	-2.0	-1.4	-3.5	0.6	-3.3
Number of IRF Days, 90-day PDP[1]		-1.3			
Number of Readmission Days, 90-day PDP[1]	-0.4	0.3	0.5	-0.2	0.9
Number of SNF Days, 90-day PDP[1]	-2.5	-1.6	-3.4	0.6	-4.5

[1] The number of days is conditional on having at least 1 day in the respective setting

Note: Statistical significance at the 0.05 level is indicated by orange and green shaded cells. Orange indicates the DiD estimate was negative and statistically significant; green indicates the DiD estimate was positive and statistically significant. The results presented in this table represent episodes initiated in the first four quarters of the BPCI Initiative (Q4 2013 through Q3 2014). A blank cell indicates that the outcome cannot be presented, either due to insufficient sample size or the type of episodes initiated during the time period. pp=percentage points. It should be noted that the columns are not mutually exclusive. The surgical sample includes all surgical episodes, including the orthopedic surgery clinical episode group. Similarly, the nonsurgical sample includes all nonsurgical episodes, including nonsurgical respiratory and nonsurgical cardiovascular. Therefore, the results of the larger surgical and nonsurgical groups are largely driven by the three clinical episode groups displayed in the table.

Source: Lewin analysis of Medicare claims and enrollment data for episodes that began Q4 2011 through Q3 2014 for BPCI providers and a comparison group.

b.   *Model 3 HHA-initiated episodes*

This section summarizes the results for the key payment outcomes for Model 3 HHA EIs relative to the comparison group.

Exhibit 103 presents the direction and point estimate of the DiD estimate for the key payment outcomes and each clinical episode group for Model 3 HHAs. It should be noted that the columns are not mutually exclusive. The nonsurgical sample includes all nonsurgical episodes, including nonsurgical respiratory and nonsurgical cardiovascular. Therefore, the results of the larger nonsurgical group are largely driven by the two clinical episode groups displayed in the table. The baseline and intervention estimates, for both BPCI and the comparison group, can be located for key outcomes in Sections IV.D through IV.I.

In general, standardized allowed payment amounts for BPCI patients were similar to patients treated by comparison providers. During the first four quarters of the initiative, the change in total standardized allowed payment for Part A and B services during the anchor hospitalization and the 90-day PDP was not statistically significant for any clinical episode group among BPCI participating Model 3 HHAs relative to the comparison group. While the change in total standardized allowed payment included in the bundle definition was negative across all clinical episode groups, these differences were not statistically significant relative to the comparison group.

During the first three quarters of the initiative, neither the change in the total amount included in the pre-bundle period, nor the change in the total amount in the post-bundle period, were statistically significant relative to the comparison group, with the exception of the statistically significant decrease in the total amount included in the pre-bundle period for the non-surgical clinical episode group (-$567).

Exhibit 103: Diff-in-Diff Estimate for Allowed Payment Outcomes, by Clinical Episode Group, Model 3 HHAs, Q4 2011-Q3 2014

Measure	Length	Surgical	Nonsurgical	Nonsurgical Respiratory	Nonsurgical Cardiovascular
Number of episodes initiated Q4 2013 – Q3 2014		471	2,761	845	1,485
BPCI Cumulative to Date (Q4 2013 – Q3 2014)					
Standardized allowed amount (Part A & B), IP through 90-day post-discharge period		-$232	-$688	-$877	$493
Standardized allowed amount in bundle definition	90	-$937	-$658	-$118	-$196
Standardized allowed not included in bundle definition	90		-$108	-$48	-$101
First Three Quarters of BPCI (Q4 2013 – Q2 2014)					
Standardized allowed amount Part A & B, Days 1-30 Post bundle payment (PBP)	90	$1,142	-$177	-$93	-$114
Standardized allowed amount Part A & B, Days 31-60 PBP	90		-$215	-$83	-$647
Standardized allowed amount Part A & B, pre-bundle period	90		-$567	-$682	$11

Note: Statistical significance at the 0.05 level is indicated by orange and green shaded cells. Orange indicates the DiD estimate was negative and statistically significant; green indicates the DiD estimate was positive and statistically significant. The results presented in this table represent episodes initiated in the first four quarters of the BPCI Initiative (Q4 2013 through Q3 2014). Please note that pre-bundle and post-bundle period (PBP) payment outcomes are reported with a one-quarter delay. Many of the payment outcomes are stratified by length of episode. A blank cell indicates that the outcome cannot be presented, either due to insufficient sample size or the type of episodes initiated during the time period. It should be noted that the columns are not mutually exclusive. The nonsurgical sample includes all nonsurgical episodes, including nonsurgical respiratory and nonsurgical cardiovascular. Therefore, the results of the larger nonsurgical group are largely driven by the two clinical episode groups displayed in the table.

Source: Lewin analysis of Medicare claims and enrollment data for episodes that began Q4 2011 through Q3 2014 for BPCI providers and a comparison group.

We calculated a number of utilization outcomes to understand the changes that contributed to the payment outcomes summarized above. See **Appendix L** for detailed outcome definitions. The baseline and intervention estimates, for both BPCI and the comparison group, can be located for key outcomes in Sections IV.D through IV.I.

As seen in Exhibit 104, in the non-surgical cardiovascular clinical episode group, there was a statistically significant decline in the number of HH visits (-2.3 visits), and in the number of days in SNF care (-3.0 days), among those who used HHA or SNF services, relative to the comparison group. However, the declines in SNF days and HH visits did not result in statistically significant payments reductions, as seen in Exhibit 103 above.

Exhibit 104: Diff-in-Diff Estimate for Utilization Outcomes, by Clinical Episode Group, Model 3 HHAs, Baseline to Intervention

Measure	Surgical	Nonsurgical	Nonsurgical Respiratory	Nonsurgical Cardiovascular
Number of episodes initiated Q4 2013 – Q3 2014	471	2,761	845	1,485
Acute Inpatient Care LOS	-0.1	-0.2	-0.2	-0.1
Number of HH Visits, 90-day PDP[1]	-0.1	-1.9	-0.3	-2.3
Number of Institutional Days, 90-day PDP[1]	1.6	-1.0	-0.2	-0.3
Number of IRF Days, 90-day PDP[1]		-1.1		
Number of Readmission Days, 90-day PDP[1]	2.1	0.2	1.3	0.3
Number of SNF Days, 90-day PDP[1]	0.4	-2.0	-1.7	-3.0

[1] The number of days/visits is conditional on having at least 1 day in the respective setting.

Note: Statistical significance at the 0.05 level is indicated by orange and green shaded cells. Orange indicates the DiD estimate was negative and statistically significant; green indicates the DiD estimate was positive and statistically significant. The results presented in this table represent episodes initiated in the first four quarters of the BPCI Initiative (Q4 2013 through Q3 2014). A blank cell indicates that the outcome cannot be presented, either due to insufficient sample size or the type of episodes initiated during the time period. pp=percentage points. It should be noted that the columns are not mutually exclusive. The nonsurgical sample includes all nonsurgical episodes, including nonsurgical respiratory and nonsurgical cardiovascular. Therefore, the results of the larger nonsurgical group are largely driven by the two clinical episode groups displayed in the table.

Source: Lewin analysis of Medicare claims and enrollment data for episodes that began Q4 2011 through Q3 2014 for BPCI providers and a comparison group.

2.   *Changes in beneficiary quality of care*

Exhibit 105 presents the impact of Model 3 BPCI on key claims-based and patient assessment-based quality of care measures by clinical episode group. The outcomes are stratified in the exhibit based on the quarters included in the DiD estimate. Claim-based outcomes were measured based on episodes initiated during the first year of the initiative, while assessment-based outcomes were measured based on episodes initiated during the first three quarters of the initiative in order to allow for sufficient assessment data run out. Orange highlighted cells represent a statistically significant decrease relative to the comparison group. Green highlighted cells represent a statistically significant increase relative to the comparison group.

The claims-based quality of care measures for SNFs and HHAs are mortality within 30 days of discharge and emergency department use and readmission rates during the first 30 days of the episode. (Hospital readmission rates and emergency department visits in the 90-day PDP were not produced for any Model 3 SNF or HHA clinical groups.) The patient assessment outcomes include improvement or change in functional status among beneficiaries who received post-acute care. See **Appendix L** for detailed outcome definitions.   The baseline and intervention estimates, for both BPCI and the comparison group, can be located for key outcomes in Sections IV.D through IV.I.

a.   *Claim-based measures*

For Model 3 SNF-initiated episodes, the quality of care for BPCI patients was similar to that of patients of non-BPCI providers, as measured through various claim-based measures, with one exception (Exhibit 105). In the non-surgical cardiovascular clinical episode group, there was a

statistically significant increase of 7.0pp in the unplanned readmission rate relative to the comparison group.

For Model 3 HHA-initiated episodes in the non-surgical cardiovascular clinical episode group, there was a statistically significant increase in the unplanned readmission rate from baseline to intervention by 4.2pp for patients of BPCI HHAs relative to the comparison providers (Exhibit 106). In the surgical clinical episode group, there was a statistically significant decrease in emergency department use by 4.4pp among patients of BPCI participating HHAs relative to the comparison group, as did the unplanned readmission rate (-6.0pp). There were no statistically significant changes in mortality, readmission rates, or emergency department visits in the non-surgical respiratory clinical episode group.

b. Assessment based measures

Overall, none of the changes in assessment-based quality measures from baseline to intervention period were statistically different for BPCI Model 3 and comparison patients, with two exceptions. There was a statistically significant negative effect of the BPCI intervention (13.9pp) on the share of patients that exhibited improvement in self-care function among BPCI Model 3 patients with orthopedic surgery episodes who were treated at a Model 3 SNF EI, compared with comparison patients. The share of patients that exhibited improvement in self-care function fell among BPCI patients but remained the same rate among the similar comparison group patients. We observed a statistically significant and positive effect of BPCI on the share of patients that exhibited improvement in upper body dressing (4.4pp) among BPCI Model 3 patients with non-surgical episodes. The share of patients that exhibited improvement increased among BPCI patients and remained unchanged among comparison patients.

Exhibit 105: Diff-in-Diff Estimate for Claim-based Quality Outcomes, by Clinical Episode Group, Model 3 SNFs, Baseline to Intervention

Measure	Surgical	Nonsurgical	Orthopedic Surgery	Nonsurgical Respiratory	Nonsurgical Cardiovascular
Number of episodes initiated Q4 2013 – Q3 2014	1,879	5,467	1,303	988	1,087
BPCI Cumulative to Date (Q4 2013 – Q3 2014)					
All-cause mortality rate, first 30 days of episode	0.6 pp	0.1 pp		-3.8 pp	1.8 pp
Emergency Department Use, first 30 days of episode	-1.6 pp	0.2 pp	-1.0 pp	-0.5 pp	0.3 pp
Unplanned Readmission Rate, first 30 days of episode	-0.5 pp	2.0 pp	1.0 pp	0.3 pp	7.0 pp
First Three Quarters of BPCI (Q4 2013 – Q2 2014)					
ADL SNF, improved mobility function	-9.1 pp	-0.6 pp	-8.0 pp	-3.7 pp	-4.2 pp
ADL SNF, improved overall function	-9.6 pp	-0.8 pp	-11.9 pp	-1.7 pp	3.0 pp
ADL SNF, improved self-care function	-7.0 pp	-1.3 pp	-13.9 pp	-3.1 pp	4.0 pp

Note: Statistical significance at the 0.05 level is indicated by orange and green shaded cells. Orange indicates the DiD estimate was negative and statistically significant; green indicates the DiD estimate was positive and statistically significant. The results presented in this table represent episodes initiated in the first four quarters of the BPCI Initiative (Q4 2013 through Q3 2014). Please note that assessment-based quality measures are reported with a one-quarter delay. A blank cell indicates that the outcome cannot be presented, either due to insufficient sample size or the type of episodes initiated during the time period. pp=percentage points. It should be noted that the columns are not mutually exclusive. The surgical sample includes all surgical episodes, including the orthopedic surgery clinical episode group. Similarly, the nonsurgical sample includes all nonsurgical episodes, including nonsurgical respiratory and nonsurgical cardiovascular. Therefore, the results of the larger surgical and nonsurgical groups are largely driven by the three clinical episode groups displayed in the table.

Source: Lewin analysis of Medicare claims and enrollment data for episodes that began Q4 2011 through Q3 2014 for BPCI providers and a comparison group.

Exhibit 106: Diff-in-Diff Estimate for Claim-based Quality Outcomes, by Clinical Episode Group, Model 3 HHAs, Baseline to Intervention

Measure	Surgical	Nonsurgical	Nonsurgical Respiratory	Nonsurgical Cardiovascular
Number of episodes initiated Q4 2013 – Q3 2014	471	2,761	845	1,485
BPCI Cumulative to Date (Q4 2013 – Q3 2014)				
All-cause mortality rate, first 30 days of episode		-1.1 pp	-2.2 pp	-0.5 pp
Emergency Department Use, first 30 days of episode	-4.4 pp	-0.5 pp	-1.6 pp	0.3 pp
Unplanned Readmission Rate, first 30 days of episode	-6.0 pp	0.4 pp	-2.1 pp	4.2 pp
First Three Quarters of BPCI (Q4 2013 – Q2 2014)				
ADL HHA, improved ambulation	4.6 pp	0.7 pp	-0.8 pp	-1.7 pp
ADL HHA, improved bathing	7.4 pp	1.2 pp	3.0 pp	-0.7 9p
ADL HHA, improved bed transferring	-1.3 pp	4.0 pp	1.6 pp	3.5 pp
ADL HHA, improved lower body dressing	2.5 pp	2.3 pp	0.4 pp	1.7 pp
ADL HHA, improved upper body dressing	2.1 pp	4.4 pp	4.4 pp	3.5 pp

Note: Statistical significance is indicated by orange and green shaded cells. Orange indicates the DiD estimate was negative and statistically significant; green indicates the DiD estimate was positive and statistically significant. The results presented in this table represent episodes initiated in the first four quarters of the BPCI Initiative (Q4 2013 through Q3 2014). Please note that assessment-based quality measures are reported with a one-quarter delay. A blank cell indicates that the outcome cannot be presented, either due to insufficient sample size or the type of episodes initiated during the time period. pp=percentage points. It should be noted that the columns are not mutually exclusive. The nonsurgical sample includes all nonsurgical episodes, including nonsurgical respiratory and nonsurgical cardiovascular. Therefore, the results of the larger nonsurgical group are largely driven by the two clinical episode groups displayed in the table.

Source: Lewin analysis of Medicare claims and enrollment data for episodes that began Q4 2011 through Q3 2014 for BPCI providers and a comparison group

c. Beneficiary survey

Exhibits 107 and 108 below report the survey results for Model 3 respondents. Response rates for all Model 3 strata can be found in **Appendix O**. BPCI respondents were less likely than comparison respondents to report improvement in planning regular tasks, a difference of 6.2pp (statistically significant at 0.05). This result was primarily driven by BPCI respondents with an MCC, among whom the difference was 8.4pp. There were no other statistically significant differences in rates of improvement or decline among the other functional measures.

Exhibit 107: Improvement and Decline in Functional Status for BPCI and Comparison Survey Respondents, Model 3, May, June, October, and November 2014

Functional Measure	Survey Strata[a]	BPCI Rate (n)	Comparison Rate (n)	Treatment Effect
Improvement in bathing, dressing, using the toilet, or eating	Overall (May, June)	53.7% (585)	56.2% (852)	-2.5 [-7.6, 2.5]
	MCC	49.6% (303)	51.7% (331)	-2.1 [-9.6, 5.5]
	Non-MCC	59.1% (282)	61.0% (521)	-1.9 [-8.5, 4.6]
	Overall (October, November)	49.9% (467)	55.1% (466)	-5.2 [-10.9, 0.5]

Functional Measure	Survey Strata[a]	BPCI Rate (n)	Comparison Rate (n)	Treatment Effect
Improvement in walking without rest	Overall (May, June)	68.1% (594)	69.2% (863)	-1.2 [-5.9, 3.6]
	MCC	64.0% (310)	65.9% (337)	-2.0 [-9.1, 5.2]
	Non-MCC	72.9% (284)	73.4% (526)	-0.5 [-6.5, 5.5]
	Overall (October, November)	24.4% (464)	24.6% (461)	-0.2 [-5.4, 4.9]
Improvement in use of mobility device (i.e., less frequent)	Overall (May, June)	72.9% (597)	72.7% (864)	0.2 [-4.3, 4.7]
	MCC	69.2% (311)	69.0% (336)	0.2 [-6.6, 7.1]
	Non-MCC	78.4% (286)	76.5% (528)	1.9 [-3.4, 7.3]
	Overall (October, November)	28.0% (471)	28.7% (467)	-0.7 [-5.6,4.1]
Improvement in using stairs	Overall (May, June)	35.7% (603)	34.6% (859)	1.1 [-3.4, 5.5]
	MCC	32.5% (315)	31.7% (335)	0.9 [-5.6, 7.3]
	Non-MCC	39.4% (288)	38.1% (524)	1.3 [-4.6, 7.1]
	Overall (October, November)	24.9% (451)	28.1% (440)	-3.2 [-8.6, 2.2]
Improvement in planning regular tasks	Overall (May, June)	40.8% (588)	47.0% (846)	-6.2* [-10.9, -1.5]
	MCC	34.9% (305)	43.3% (327)	-8.4* [-15.2, -1.5]
	Non-MCC	49.5% (283)	50.8% (519)	-1.4 [-7.6, 4.9]
	Overall (October, November)	36.3% (471)	39.6% (473)	-3.3 [-8.7, 2.0]
Improvement in physical/emotional problems limiting social activities (i.e., less frequent)[b]	Overall (October, November)	44.33% (466)	48.69% (461)	-4.4 [-10.7, 2.0]
Improvement in pain limiting regular activities (i.e., less frequent)[b]	Overall (October, November)	47.37% (467)	51.37% (459)	-4.0 [-10.1, 2.1]
Decline in bathing, dressing, using the toilet, or eating	Overall (May, June)	27.8% (585)	25.1% (852)	2.7 [-2.0, 7.4]
	MCC	31.1% (303)	30.6% (331)	0.6 [-6.5, 7.6]
	Non-MCC	23.4% (282)	19.5% (521)	3.9 [-2.1, 9.8]
	Overall (October, November)	26.9% (467)	25.0% (466)	1.9 [-3.3, 7.1]

Functional Measure	Survey Strata[a]	BPCI Rate (n)	Comparison Rate (n)	Treatment Effect
Decline in walking without rest	Overall (May, June)	19.7% (493)	16.4% (731)	3.3 [-1.5, 8.1]
	MCC	22.6% (251)	20.0% (275)	2.6 [-5.0, 10.1]
	Non-MCC	17.0% (242)	12.3% (456)	4.7 [-1.1, 10.6]
	Overall (October, November)	43.3% (464)	44.1% (461)	-0.8 [-6.5, 4.9]
Decline in use of mobility device (i.e., more frequent)	Overall (May, June)	18.9% (597)	19.8% (864)	-0.8 [-5.0, 3.3]
	MCC	21.7% (311)	23.3% (336)	-1.6 [-7.9, 4.7]
	Non-MCC	15.2% (286)	16.0% (528)	-0.8 [-5.7, 4.0]
	Overall (October, November)	56.4% (471)	53.6% (467)	2.8 [-2.5, 8.1]
Decline in using stairs	Overall (May, June)	55.8% (603)	57.0% (859)	-1.2 [-5.6, 3.2]
	MCC	57.6% (315)	58.3% (335)	-0.7 [-7.2, 5.8]
	Non-MCC	53.4% (288)	55.5% (524)	-2.0 [-7.7, 3.6]
	Overall (October, November)	54.5% (451)	50.7% (440)	3.8 [-1.6, 9.2]
Decline in planning regular tasks	Overall (May, June)	25.4% (448)	23.8% (674)	1.6 [-3.8, 6.9]
	MCC	30.0% (224)	26.8% (244)	3.1 [-5.5, 11.8]
	Non-MCC	19.5% (224)	20.6% (430)	-1.1 [-7.4, 5.1]
	Overall (October, November)	36.8% (471)	35.6% (473)	1.2 [-4.4, 6.8]
Decline in physical/emotional problems limiting social activities (i.e., more frequent)[b]	Overall (October, November)	32.9% (466)	30.5% (461)	2.4 [-3.7, 8.5]
Decline in pain limiting regular activities (i.e., more frequent)[b]	Overall (October, November)	23.4% (467)	24.1% (459)	-0.7 [-6.2, 4.8]

* $p < 0.05$;

Notes: Treatment effect reported in percentage points. 95% confidence interval is in brackets. Survey results were weighted to adjust for sampling and nonresponse.

MCC refers to episodes with major complicating conditions; non-MCC refers to episodes without a major complicating condition.

The questions and possible responses for walking, use of stairs, and use of a mobility device changed drastically between Waves 1 and 2. Therefore, rates of improvement and decline for these measures may differ substantially between May/June and October/November.

BPCI and comparison samples sizes for each question are reported in parentheses. The sample sizes vary between questions due to varying rates of item-response.

[a] The "overall" stratum was sampled in May/June 2014 and again in October/November 2014. Results for the first overall stratum refer to summer 2014, while the second overall stratum refers to fall 2014.

[b] Rates of change in these measures can only be estimated for the fall 2014 sample.

As in Model 2, there were few statistically significant differences between BPCI and comparison respondents in perceived health care experience. However, Exhibit 108 shows that BPCI respondents in fall 2014 were 3.7pp less likely to agree or strongly agree that they had a good understanding of how to take care of themselves before returning home, although nearly 92% of BPCI respondents did still agree or strongly agree.

Exhibit 108: Health Care Experience for BPCI and Comparison Survey Respondents, Model 3, May, June, October, November 2014

Health Care Experience Measures	Survey Strata[a]	BPCI Rate (n)	Comparison Rate (n)	Treatment Effect
Thinking about all the care you received in the hospital before and afterwards, how often did you, your family, or your caregiver get conflicting advice from medical staff about your treatment? 1[response = Never]	Overall (May, June)	56.9% (616)	60.5% (875)	-3.6 [-8.9, 1.7]
	MCC	54.3% (317)	57.8% (337)	-3.6 [-11.3, 4.2]
	non-MCC	61.2% (299)	63.0% (538)	-1.8 [-8.7, 5.1]
	Overall (October, November)	62.1% (473)	66.5% (479)	-4.4 [-10.9, 2.2]
Thinking about all of the care you received in the hospital and afterwards, how often were the services you got appropriate for the level of care you needed? 1[response = Always]	Overall (May, June)	48.2% (606)	47.6% (871)	0.7 [-4.7, 6.1]
	MCC	46.2% (312)	47.6% (339)	-1.4 [-9.2, 6.4]
	non-MCC	51.4% (294)	47.5% (532)	3.9 [-3.4, 11.3]
	Overall (October, November)	48.8% (480)	51.0% (477)	-2.2 [-8.9, 4.5]
Thinking about when you left the hospital, were you discharged at the right time? 1[response = Yes]	Overall (May, June)	82.7% (606)	84.8% (871)	-2.1 [-6.2, 2.1]
	MCC	79.1% (315)	83.9% (339)	-4.8 [-11.0, 1.4]
	non-MCC	88.6% (291)	85.6% (532)	3.1 [-2.0, 8.2]
	Overall (October, November)	85.8% (478)	85.1% (477)	0.7 [-4.4, 5.8]
Thinking about when you left the hospital listed in the cover letter, the medical staff took your preferences and those of your family or your caregiver into account in deciding what health care services you should have after you left the hospital. 1[response = Agree/Strongly Agree]	Overall (May, June)	90.3% (546)	89.6% (794)	0.7 [-2.8, 4.2]
	MCC	90.1% (281)	89.7% (309)	0.3 [-4.8, 5.5]
	non-MCC	90.7% (265)	89.5% (485)	1.2 [-3.6, 5.9]
	Overall (October, November)	91.3% (431)	90.7% (440)	0.7 [-3.3, 4.6]

Health Care Experience Measures	Survey Strata[a]	BPCI Rate (n)	Comparison Rate (n)	Treatment Effect
Before you prepared to go home (or to someone else's home, or to an assisted living facility), you and your family or caregiver had a good understanding of how to take care of yourself. 1[response = Agree/Strongly Agree]	Overall (May, June)	92.0% (519)	92.4% (723)	-0.4 [-3.7, 2.9]
	MCC	93.4% (257)	92.8% (271)	0.6 [-4.0, 5.2]
	non-MCC	90.0% (262)	92.1% (452)	-2.1 [-7.0, 2.8]
	Overall (October, November)	91.5% (419)	95.3% (420)	-3.7* [-7.4, -0.1]
Before you prepared to go home (or to someone else's home, or to an assisted living facility), medical staff clearly explained how to take your medications. 1[response = Agree/Strongly Agree]	Overall (May, June)	89.5% (500)	88.6% (679)	0.9 [-3.1, 4.9]
	MCC	88.3% (256)	88.4% (261)	-0.1 [-5.9, 5.8]
	non-MCC	91.5% (244)	88.8% (418)	2.6 [-2.4, 7.7]
	Overall (October, November)	90.5% (417)	89.8% (409)	0.7 [-3.5, 5.0]
Before you prepared to go home (or to someone else's home, or to an assisted living facility), medical staff clearly explained what follow-up appointments or treatments would be needed 1[response = Agree/Strongly Agree]	Overall (May, June)	92.0% (517)	90.7% (724)	1.3 [-2.2, 4.8]
	MCC	90.6% (261)	91.3% (278)	-0.8 [-6.0, 4.5]
	non-MCC	94.3% (256)	90.1% (446)	4.2 [-0.1, 8.4]
	Overall (October, November)	92.2% (394)	93.8% (394)	-1.6 [-5.2, 2.1]
Overall, since you returned home (or to someone else's home, or to an assisted living facility), you and your caregivers have been able to manage your health needs. 1[response = Agree/Strongly Agree]	Overall (May, June)	92.6% (531)	94.4% (754)	-1.8 [-5.0, 1.3]
	MCC	92.9% (267)	94.8% (291)	-1.9 [-6.3, 2.5]
	non-MCC	92.2% (264)	94.1% (463)	-1.9 [-6.3, 2.5]
	Overall (October, November)	96.7% (409)	97.2% (405)	-0.5 [-2.9, 1.9]
Overall, how satisfied are you with your recovery since you left the hospital? 1[response = Extremely/Quite a Bit]	Overall (May, June)	55.0% (601)	58.1% (860)	-3.1 [-8.5, 2.3]
	MCC	52.9% (314)	57.8% (337)	-4.9 [-12.7, 2.8]
	non-MCC	58.4% (287)	58.4% (523)	0.1 [-7.3, 7.4]
	Overall (October, November)	59.1% (456)	62.0% (459)	-3.0 [-9.7, 3.8]

Notes: * p < 0.05. Results are unadjusted. Treatment effect reported in percentage points. 95% confidence interval is in brackets. Survey results were weighted to adjust for sampling and nonresponse.

BPCI and comparison samples sizes for each question are reported in parentheses. The sample sizes vary between questions due to varying rates of item-response.

[a] The "overall" stratum was sampled in May/June 2014 and again in October/November 2014. Results for the first overall stratum refer to summer 2014, while the second overall stratum refers to fall 2014.

Similar to Model 2, there were few statistically significant differences in health status or health care experience outcomes between BPCI Model 3 respondents and comparison respondents. However, every statistically significant point estimate indicates an inferior outcome among BPCI Model 3 respondents. This may suggest declines in the quality of care among BPCI providers in Model 3, though we note that no statistically significant results were persistent from summer 2014 through the fall. Additional survey data will help to clarify whether care quality is trending worse or if these are simply isolated results.

We do not report results for Model 3 respondents at the clinical episode group level as we did not survey BPCI model 3 and comparison patients separately by clinical episode groups in the summer or fall.

C. Market Dynamics

We hypothesized that Model 3 EIs may seek to attract patients from hospitals with which they have a good working relationship or that provide better quality of care and better transition planning. We also hypothesized that the hospital EIs might strive to increase their market share of certain BPCI-eligible episodes in order to increase efficiency and savings. Two PAC-from-hospitals concentration indices were created to examine whether the BPCI program contributed to a change in the shares of patients discharged from hospitals and admitted to Model 3 PAC EIs. We also calculated the market share of Model 3 PAC EIs to determine whether EIs have captured a greater share of BPCI-eligible PAC episodes over time. For each SNF EI, the SNF-from-hospitals concentration index is calculated as the proportion of patients that transitioned from each hospital to the SNF, with no other PAC stay in between, squared and summed across all hospitals. The index ranges from zero to one, where a higher index indicates greater shares of patients admitted to the SNF EI coming from fewer hospitals. We calculated the same index for patients who transitioned from hospitals to HHA EIs, with no other PAC stay in between. The numerator for calculating the market share of a SNF EI is the number of BPCI-eligible episodes admitted to the SNF within 30 days of any hospital discharge, and with no other SNF stay in between. The denominator is the total number of such patients admitted, summed over all SNFs in the EI's market. The market share for HHAs was calculated in the same manner.

We calculated these measures separately for MJRLE, CHF, and sepsis for a two-year pre-BPCI period and a one year intervention period. Since MJRLE has the largest volume of all clinical episode groups and is currently a high-priority area for CMS, we focus on these episodes and report results for CHF and sepsis episodes in **Appendix P**. CBSAs were used to define a Model 3 EI's market.

The discussion below is divided into two sub-sections. The first sub-section presents trends in the SNF-from-hospitals and HHA-from-hospitals concentration indices for MJRLE episodes. The second sub-section presents the market share of MJRLE episodes for SNF EIs and HHA EIs separately. Average rates are also reported for selected individual markets. For Model 3 SNFs we report Chicago-Naperville-Joliet, IL (Chicago market) and Warren-Farmington-Hills-Troy, MI (Suburban Detroit market). For Model 3 HHAs we report Charleston-North Charleston, SC (Charleston market) and Jacksonville, FL (Jacksonville market). Of note, given the small counts of EIs and the preliminary nature of these analyses, some market effects may have gone undetected.

1. PAC-from-hospitals concentration index for Model 3 EIs

Exhibits 109 and 110 present the descriptive trends for the SNF-from-hospitals and HHA-from-hospitals concentration indices for MJRLE episodes over all Model 3 EIs. The average highest percent of admissions to each EI from one hospital and the average number of inpatient hospitals whose patients received care at each EI are also shown. Below the all-market results are the average concentration indices for MJRLE episodes across Model 3 SNF EIs in the Chicago and Suburban Detroit markets (Exhibit 109) and across Model 3 HHA EIs in the Charleston and Jacksonville markets (Exhibit 110). For context, the tables also show the total number of hospitals in the individual markets with BPCI-eligible MJRLE episodes admitted to SNFs or HHAs. [66]

The average SNF-from-hospitals and HHA-from-hospitals concentration indices for MJRLE episodes did not change over time.[67] For both SNF and HHA Model 3 EIs, the average number of hospitals whose patients transitioned to a Model 3 EI also remained the same over time, as did the average highest percent of MJRLE patients transitioning from one hospital to a Model 3 EI.

There was little change in the average concentration of MJRLE patients admitted to SNF EIs in the Chicago market after a hospital discharge. However, the average SNF-from-hospitals concentration index for MJRLE episodes in the suburban Detroit market was highest in the Q2 2014/Q3 2014 intervention period (0.41) compared to those in the baseline periods (0.27-0.33). Only two HHA EIs were located in the Charleston market, and we saw no change in their average concentration of MJRLE patients being admitted from certain hospitals. There was only one HHA EI located in the Jacksonville market and its HHA-from-hospitals index for MJRLE episodes was 0.68 and 0.64 in the two BPCI periods compared to the next highest index of 0.58 in the pre-BPCI period.

Appendix P shows the trends for CHF and sepsis episodes. We saw no indication that the average SNF-from-hospitals or HHA-from-hospitals concentration indices over all Model 3 EIs nationwide had changed between the pre-BPCI and BPCI periods, for CHF or sepsis episode types.

---

[66] Total number of hospitals in a market includes only those with at least one admission of a patient in the particular clinical episode group who was then discharged to a SNF or HHA during a six-month period. The hospitals can locate in a different market than the market where the Model 3 EIs are located. For example, a Model 3 EI can admit patients who were treated at a geographically close hospital located in a neighboring CBSA. This number varies by clinical episode group and time period.

[67] We did not estimate statistical significance of the difference in measure rates between the baseline and intervention period due to small sample size. Even at the national level (all EIs across all markets), the power analysis suggested the sample size was too small to reliably detect a statistically significant difference if one was present. We will re-evaluate this next year.

Exhibit 109: Average SNF-from-hospitals Concentration Index for Major Joint Replacement of the Lower Extremity Episodes, Model 3, Q2 2011-Q3 2014

	Baseline				Intervention	
	Q4 2011 / Q1 2012	Q2 2012 / Q3 2012	Q4 2012 / Q1 2013	Q2 2013 / Q3 2013	Q4 2013 / Q1 2014*	Q2 2014 / Q3 2014
ALL MARKETS (51 SNF EIs)						
SNF-from-hospitals concentration index						
Mean	0.44	0.43	0.41	0.43	0.43	0.44
Median	0.38	0.38	0.35	0.38	0.39	0.38
25th Pctl	0.30	0.25	0.25	0.25	0.25	0.30
75th Pctl	0.50	0.56	0.50	0.50	0.56	0.54
Number of hospitals patients were received from, per Model 3 EI						
Mean	4.2	4.4	4.7	4.3	4.8	4.3
Highest percent of patients from one hospital						
Mean	55.1%	53.8%	50.0%	51.2%	53.0%	53.5%
CHICAGO MARKET (10 SNF EIs)						
Mean concentration	0.49	0.45	0.48	0.44	0.51	0.41
SD	0.28	0.22	0.31	0.24	0.27	0.24
Number of hospitals patients were received from, per SNF EI						
Mean	5.2	4.9	5.7	5.7	5.3	4.9
Highest percent of patients from one hospital						
Mean	60.2%	56.8%	57.0%	54.7%	62.0%	53.6%
Number of hospitals in market with patients admitted to SNFs						
Total	72	72	77	78	72	76
SUBURBAN DETROIT MARKET (4 SNF EIs)						
Mean concentration	0.33	0.32	0.27	0.28	0.24	0.41
SD	0.07	0.07	0.06	0.06	0.05	0.18
Number of hospitals patients were received from, per SNF EI						
Mean	5.8	4.5	6.8	6.0	6.8	5.3
Highest percent of patients from one hospital						
Mean	48.1%	41.7%	39.4%	38.5%	31.8%	54.5%
Number of hospitals in market with patients admitted to SNFs						
Total	32	36	33	35	32	33

Note: Rates were calculated based on 51 Model 3 SNF EIs with patients admitted from hospitals over all six periods.
* A majority of Model 3 SNF EIs joined BPCI in Q1 2014. *Chicago and Suburban Detroit*: All Model 3 SNF EIs joined BPCI in Q1 2014.

Source: Lewin analysis of Medicare claims and enrollment data for episodes that began Q4 2011 through Q3 2014 for BPCI participants.

Exhibit 110: Average HHA-from-hospitals Concentration Index for Major Joint Replacement of the Lower Extremity Episodes, Model 3, Q2 2011-Q3 2014

	Baseline				Intervention	
	Q4 2011 / Q1 2012	Q2 2012 / Q3 2012	Q4 2012 / Q1 2013	Q2 2013 / Q3 2013	Q4 2013 / Q1 2014*	Q2 2014 / Q3 2014
ALL MARKETS (26 HHA EIs)						
HHA-from-hospitals concentration index						
Mean	0.30	0.35	0.33	0.35	0.40	0.34
Median	0.28	0.33	0.32	0.30	0.32	0.29
25th Pctl	0.21	0.23	0.25	0.23	0.24	0.21
75th Pctl	0.39	0.46	0.40	0.41	0.51	0.41
Number of hospitals patients were received from, per Model 3 EI						
Mean	8.7	7.7	8.7	8.5	9.0	8.9
Highest percent of patients from one hospital						
Mean	42.7%	48.8%	47.9%	47.7%	52.7%	47.2%
CHARLESTON MARKET (2 HHA EIs)						
Mean concentration	0.40	0.35	0.30	0.28	0.31	0.31
SD	0.02	0.02	0.02	0.01	0.03	0.07
Number of hospitals patients were received from, per HHA EI						
Mean	6.0	5.5	6.5	6.0	6.0	4.5
Highest percent of patients from one hospital						
Mean	54.1%	47.8%	41.3%	38.1%	41.9%	38.8%
Number of hospitals in market with patients admitted to HHAs						
Total	11	11	13	12	12	13
JACKSONVILLE MARKET (1 HHA EI)						
Mean concentration	0.58	0.56	0.41	0.45	0.68	0.64
Number of hospitals patients were received from, per HHA EI						
Mean	4	4	5	4	7	5
Highest percent of patients from one hospital						
Mean	64.9%	50.9%	42.3%	27.5%	41.7%	42.9%
Number of hospitals in market with patients admitted to HHAs						
Total	12	12	13	14	14	13

Note: Rates were calculated based on 26 Model 3 HHA EIs with patients admitted from hospitals over all six periods.
* A majority of Model 3 HHA EIs joined BPCI in Q1 2014. *Charleston:* All Model 3 HHA EIs joined BPCI in Q1 2014. *Jacksonville*: All Model 3 HHA EIs joined BPCI in Q4 2013.
Source: Lewin analysis of Medicare claims and enrollment data for episodes that began Q4 2011 through Q3 2014 for BPCI participants.

2. *Market share of Model 3 EIs*

Exhibits 111 and 112 present the descriptive trends for the market shares of MJRLE episodes across all Model 3 SNF EIs and Model 3 HHA EIs, and in the selected individual markets. The average number of MJRLE episodes per SNF or HHA EI is also shown. The average market share of MJRLE episodes across all SNF EIs was around 4.5% and changed little in the BPCI period.

Across all HHA EIs, the average market share of MJRLE episodes increased from 8.5% before BPCI to 9.9% in the BPCI period. This upward trend appears to have started before the beginning of the BPCI program, however, so we cannot conclude that the change in market share is attributed the BPCI program. The median market shares across all SNF EIs and HHA EIs were lower than the average market shares and showed little or no change over time.

Turning to the individual markets, the market shares of MJRLE episodes for SNF EIs in the Chicago and Suburban Detroit markets remained steady over time, at about 1% in the Chicago market and 3% in the Suburban Detroit market.[68] The average market share of MJRLE episodes for the two HHA EIs in the Charleston market remained steady at around 14%. However, the single HHA EI in the Jacksonville market saw its market share increase from a range of 12.7% to 18.8% in the pre-BPCI periods, to 22% and 20.3% in the two BPCI intervention periods.[69] This is the only result under Model 3 that is consistent with our hypothesis that BPCI EIs will attempt to increase their market share of BPCI episodes to take advantage of efficiencies and savings. The time trends for the market shares of CHF and sepsis episodes are presented in **Appendix P**. Briefly, we see no evidence of statistically significant changes between the pre-BPCI and one year intervention periods.

---

[68] Given the large size of these markets and the small size of the EIs' market shares, it will be difficult to detect any significant changes in the EIs market shares regardless of the magnitude of the change. Hence, for the next annual report, we will explore alternative metrics to capture substantive changes in EI market shares in large markets like these.

[69] This result is for a single EI but consistent with our hypothesis that BPCI EIs will attempt to increase their market share of BPCI episodes to take advantage of efficiencies and savings. Hence, for the next annual report, we will explore alternative approaches to better examine changes happening at the individual EI levels that are masked by statistics that aggregate the EIs at the national or individual market levels.

Exhibit 111: Average Market Share of Episode Initiators for Major Joint Replacement of the Lower Extremity Episodes, Model 3 SNFs, Q2 2011-Q3 2014

	Baseline				Intervention	
	Q4 2011 / Q1 2012	Q2 2012 / Q3 2012	Q4 2012 / Q1 2013	Q2 2013 / Q3 2013	Q4 2013 / Q1 2014*	Q2 2014 / Q3 2014
ALL MARKETS (52 SNF EIs)						
SNF EI Market Share						
Mean	4.6%	4.8%	4.4%	4.3%	4.6%	4.3%
SD	6.7%	7.4%	6.9%	7.1%	6.6%	6.6%
Median	1.5%	1.3%	1.3%	1.6%	2.1%	1.4%
25th Pctl	0.7%	0.5%	0.5%	0.5%	0.6%	0.5%
75th Pctl	5.7%	5.7%	5.2%	4.3%	4.4%	4.7%
Number of MJRLE episodes per SNF EI						
Mean	18.8	22.9	22.8	22.2	23.2	21.0
CHICAGO MARKET (10 SNF EIs)						
SNF EI Market Share						
Mean	0.9%	0.9%	1.0%	0.9%	0.8%	0.7%
SD	1.1%	1.2%	1.4%	1.3%	1.0%	0.8%
Number of MJRLE episodes per SNF EI						
Mean	25.2	26.9	28.1	28.0	24.2	22.0
Number of SNFs with MJRLE admissions in the market						
Total	207	217	209	205	217	216
SUBURBAN DETROIT MARKET (4 SNF EIs)						
SNF EI Market Share						
Mean	2.6%	2.5%	3.6%	2.8%	3.2%	2.8%
SD	2.7%	3.3%	4.0%	3.3%	3.9%	3.4%
Number of MJRLE episodes per SNF EI						
Mean	19.0	26.0	35.3	29.0	30.8	26.3
Number of SNFs with MJRLE admissions in the market						
Total	56	63	63	61	67	61

* A majority of Model 3 SNF EIs joined BPCI in Q1 2014. Chicago and Suburban Detroit: All Model 3 SNF EIs joined BPCI in Q1 2014.

SD = standard deviation.

Source: Lewin analysis of Medicare market share data from Q4 2011-Q3 2014

Exhibit 112: Average Market Share of Episode Initiators for Major Joint Replacement of the Lower Extremity Episodes, Model 3 HHAs, Q2 2011-Q3 2014

	Baseline				Intervention	
	Q4 2011 / Q1 2012	Q2 2012 / Q3 2012	Q4 2012 / Q1 2013	Q2 2013 / Q3 2013	Q4 2013 / Q1 2014*	Q2 2014 / Q3 2014
ALL MARKETS (27 HHA EIs)						
HHA EI Market Share						
Mean	9.0%	8.5%	9.0%	9.4%	9.9%	9.9%
SD	10.3%	10.1%	10.0%	10.0%	10.1%	9.9%
Median	6.3%	5.4%	5.5%	6.3%	7.6%	6.6%
25th Pctl	2.8%	2.8%	2.8%	2.7%	3.0%	3.4%
75th Pctl	10.8%	11.9%	12.9%	12.3%	13.3%	13.7%
Number of MJRLE episodes per HHA EI						
Mean	79.0	77.3	81.7	85.5	95.0	84.3
CHARLESTON MARKET (2 HHA EIs)						
HHA EI Market Share						
Mean	13.0%	14.0%	15.3%	14.4%	13.3%	14.0%
SD	2.8%	0.6%	1.2%	3.5%	7.7%	9.4%
Number of MJRLE episodes per HHA EI						
Mean	52.0	51.5	61.5	64.0	59.0	59.0
Number of HHAs with MJRLE admissions in the market						
Total	7	7	8	8	8	9
JACKSONVILLE MARKET (1 HHA EI)						
HHA EI Market Share						
Mean	10.5%	12.7%	18.8%	17.4%	22.0%	20.3%
Number of MJRLE episodes per HHA EI						
Mean	88	92	144	118	152	107
Number of HHAs with MJRLE admissions in the market						
Total	35	38	31	36	33	31

* A majority of Model 3 HHA EIs joined BPCI in Q1 2014. *Charleston*: All Model 3 HHA EIs joined BPCI in Q1 2014. *Jacksonville*: All Model 3 HHA EIs joined BPCI in Q4 2013.

SD = standard deviation.

Source: Lewin analysis of Medicare market share data from Q4 2011-Q3 2014

D.   Impact of BPCI on SNF Orthopedic Surgery

1.   Participants

During the first four quarters of the initiative, 58 SNFs, or 92% of SNF EIs in Model 3, participated in at least one orthopedic surgery episode. There were 1,303 orthopedic surgery episodes initiated in Model 3-participating SNFs, accounting for 18% of all Model 3 SNF episodes. Approximately 76% of these episodes were for major joint replacement of the lower extremity MS-DRGs. Each of the SNFs participating in BPCI with these clinical episodes was an urban facility and 19% were non-profits. They were more likely than the typical SNF to have an inpatient rehabilitation facility

in their core based statistical area (52% vs. 29%), suggesting more post-acute care options are available than in non-participating SNF markets.

2. Patient population characteristics

Beneficiaries who were treated in a SNF participating in Model 3 for an orthopedic surgery episode were similar to all Medicare beneficiaries treated in the same MS-DRG with respect to gender and proportion eligible for Medicare due to disability (see Exhibit 113). The major differences between the BPCI and non-BPCI groups were related to age and dual eligibility status. The patients of BPCI providers had a larger proportion aged 65-79 (55.8% vs. 44.7%) and a smaller proportion aged 80 and older (38.6% vs. 46.9%) than all beneficiaries with the same MS-DRG. Among BPCI patients, 9.3% were dual eligible, compared to 18.7% of all Medicare beneficiaries.

Exhibit 113: Characteristics of BPCI Beneficiaries and All Medicare Beneficiaries with a SNF Admission for an Orthopedic Surgery MS-DRG, Model 3, Q4 2013 - Q3 2014

Characteristics	Model 3 BPCI SNF Beneficiaries with Orthopedic Surgery Episodes, Q4 2013-Q3 2014 (N = 1,187)		All Medicare Beneficiaries With SNF Admission following inpatient admission with a Orthopedic Surgery MS-DRG, Q3 2014 (N = 89,776)	
	N	%	N	%
Age				
20-64	67	5.6%	7,607	8.5%
65-79	662	55.8%	40,086	44.7%
80+	458	38.6%	42,083	46.9%
Gender				
Female	862	72.6%	64,111	71.4%
Male	325	27.4%	25,665	28.6%
Medicaid and Disability				
% Eligible Medicaid	110	9.3%	16,799	18.7%
% Disability, no ESRD	100	8.4%	8,867	9.9%

Source: Lewin analysis of Medicare claims and enrollment data for SNF-initiated episodes that began Q4 2011 through Q3 2014 for BPCI providers and all admissions to non-participating SNFs that had a preceding inpatient stay for the same MS-DRG in Q3 2014.

3. Change in Medicare standardized allowed payments

There were no statistically significant differences between BPCI episodes and episodes among comparison providers in the change from the baseline to the intervention periods in Medicare standardized allowed payments. The average total standardized allowed payment included in the bundle definition for 60- and 90-day episodes for patients treated by BPCI providers declined from baseline to intervention periods, but this reduction was not statistically significant relative to the change for patients treated by matched comparison providers (see Exhibit 114). The decline in the total included in the bundle was not statistically significant relative to the comparison group despite a statistically significant decrease in SNF payments during the 90 days post qualifying hospital discharge relative to the comparison group (- $2,101). The SNF allowed amount decreased from $11,999 to $10,194 for baseline to intervention for BPCI patients

compared to an increase for the comparison group from $11,301 to $11,597. The decrease in SNF allowed amount was driven by a statistically significant decline in the number of SNF days. The number of SNF days decreased from 21.2 to 19.3 for baseline to intervention for BPCI patients compared to an increase for the comparison group from 21.9 to 23.5, resulting in a relative decrease of 3.4 days. There were no indications that these results were achieved by shifting services outside of the bundle definition or bundle period.

Exhibit 114: Impact of BPCI on Total Standardized Allowed Payment Included in the Bundle Definition for Orthopedic Surgery Episodes, Model 3 SNFs, Q4 2013 - Q3 2014

Measure	Bundle Length	BPCI		Comparison		Diff-in-Diff estimate		
		Baseline	Intervention	Baseline	Intervention	Diff-in-Diff	LCI	UCI
Total Standardized Allowed Payment Included In The Bundle Definition *(BPCI N= 645; Comparison N= 644)*	60	$12,749	$12,623	$13,137	$14,068	-$1,058	-$2,948	$832
Total standardized allowed payment included in the bundle definition *(BPCI N= 538; Comparison N= 539)*	90	$21,115	$19,760	$20,405	$20,336	-$1,285	-$4,605	$2,034

* Denotes statistical significance at the 5% level.

LCI = lower confidence interval at the 5% level; UCI = upper confidence interval at the 5% level.

Note: Baseline is defined as episodes that began Q4 2011 through Q3 2012. Intervention is defined as episodes that began Q4 2013 through Q3 2014. Sample sizes reflect number of episodes initiated during the intervention period that met inclusion criteria for the outcome.

Source: Lewin analysis of Medicare claims and enrollment data for episodes that began Q4 2011 through Q3 2014 for BPCI and comparison providers.

4.   Change in beneficiary quality of care

The change in hospital readmission rates and emergency department visits following the start of the episode for SNF orthopedic surgery episodes was not statistically different for patients of BPCI providers and comparison providers (see Exhibit 115).

Exhibit 115: Impact of BPCI on Claims-based Quality Measures for Orthopedic Surgery Episodes, Model 3 SNFs, Q4 2013 - Q3 2014

Measure	BPCI (N=1,182)		Comparison (N=1,180)		Diff-in-Diff estimate		
	Baseline	Intervention	Baseline	Intervention	Diff-in-Diff	LCI	UCI
Unplanned Readmission Rate, within 30 days of episode start	7.0%	8.1%	7.1%	7.2%	1.0	-2.5	4.6
Emergency Department Visit Rate, within 30 days of episode start	6.0%	4.7%	5.5%	5.1%	-1.0	-3.3	1.4
Mortality, within 30 days of episode start**	n/a	n/a	n/a	n/a	n/a	n/a	n/a

* Denotes statistical significance at the 5% level.

LCI = lower confidence interval at the 5% level; UCI = upper confidence interval at the 5% level.

**risk-adjusted results were not available for this outcome due to lack of variation

Note: Baseline is defined as episodes that began Q4 2011 through Q3 2012. Intervention is defined as episodes that began Q4 2013 through Q3 2014. Sample sizes reflect number of episodes initiated during the intervention period that met inclusion criteria for the outcome.

Source: Lewin analysis of Medicare claims and enrollment data for episodes that began Q4 2011 through Q3 2014 for BPCI and comparison providers.

Among patients with orthopedic surgery episodes who were treated at a SNF, the BPCI intervention was associated with a statistically significant decline of 13.9pp in the share of patients that exhibited improvement in self-care function, relative to the comparison patients. None of the other assessment based measures were statistically different between the BPCI and comparison groups. This will continue to be monitored to determine if this finding persists.

Exhibit 116: Impact of BPCI on Assessment-based Quality Measures for Orthopedic Surgery Episodes, Model 3 SNFs, Q4 2013 – Q2 2014

Measure	BPCI (N=473)		Comparison (N=659)		Diff-in-Diff estimate		
	Baseline	Intervention	Baseline	Intervention	Diff-in-Diff	LCI	UCI
Percent of SNF patients with improved self-care function	58.9%	45.1%	57.0%	57.2%	-13.9*	-27.3	-0.6

* Denotes statistical significance at the 5% level.

LCI = lower confidence interval at the 5% level; UCI = upper confidence interval at the 5% level.

Note: Baseline is defined as episodes that began Q4 2011 through Q3 2012. Intervention is defined as episodes that began Q4 2013 through Q2 2014. Sample sizes reflect number of episodes initiated during the intervention period that met inclusion criteria for the outcome.

Source: Lewin analysis of patient assessment data for episodes that began Q4 2013 through Q2 2014 for BPCI and comparison providers.

5. *Physician group practice (PGP) SNF-initiated episodes*

There was one PGP EI that participated in Model 3 orthopedic surgery episodes during the first year of the BPCI initiative.[70] We risk-adjusted the PGP results and compared them to all Model 3

---

[70] Please note that the data linking individual physicians with their PGP contained errors, so these results should be viewed with caution.

BPCI episodes in the same model and clinical episode group. This section compares the risk-adjusted payment, utilization, and quality of care outcomes of Model 3 PGP episodes initiated in a SNF to the observed mean of all Model 3 orthopedic surgery episodes initiated in a SNF during the first year of the initiative. These results should be viewed as preliminary and with caution given the difficulty in linking individual physicians with their PGP.

The Model 3 PGP EI initiated 1,052 orthopedic surgery episodes in a SNF during the first year of BPCI (approximately 45% of all Model 3 SNF orthopedic surgery episodes). The characteristics of the PGP SNF-initiated orthopedic surgery episode patients were similar to all Model 3 SNF orthopedic surgery episode patients with regards to age, gender, prior health care utilization, and average MS-DRG case weight. The PGP SNF-initiated orthopedic surgery episode patients were less likely to be eligible for Medicaid (6.0% vs. 8.3%) and less likely to be eligible for Medicare due to a disability (6.1% vs. 8.0%) than all Model 3 SNF orthopedic surgery episode patients. Additionally, PGP SNF-initiated orthopedic surgery episode patients had a lower HCC index than all Model 3 SNF orthopedic surgery episode patients (0.392 vs. 0.515), indicating that patients in PGP SNF-initiated episodes were healthier prior to their qualifying hospital stay.

The Medicare Part A and B payments within the 90-day bundle period for Model 3 PGP orthopedic surgery episodes initiated in a SNF was lower relative to all Model 3 SNF orthopedic surgery episodes, and statistically significant at 0.05 (see Exhibit 117). The lower payment for SNF services among PGP-initiated orthopedic surgery episodes ($9,385) relative to all Model 3 SNF-initiated orthopedic surgery episodes ($10,243) contributed to this difference.

Exhibit 117: Total Standardized Allowed Payment Included in the Bundle Definition for PGP-initiated Orthopedic Surgery Episodes and All SNF Orthopedic Surgery Episodes, Model 3, Q4 2013-Q3 2014

Measure	Bundle Length	PGP (N=1,017)			All Model 3 SNF episodes (N=1,591)	Model 3 PGP Risk-adjusted Mean, vs. All Model 3 SNF, Observed Mean
		Risk-adjusted mean	LCI	UCI	Observed Mean	
Total Standardized Allowed Payment Included In The Bundle	90	$15,628	$14,899	$16,357	$16,693	-$1,065*

* Denotes statistical significance at the 5% level.
LCI = lower confidence interval at the 5% level; UCI = upper confidence interval at the 5% level.
Source: Lewin analysis of Medicare claims and enrollment data for episodes that began Q4 2013 through Q3 2014 for BPCI PGP and hospital providers.

The quality of care for BPCI patients with PGP orthopedic surgery episodes initiated in a SNF was similar to that of all Model 3 SNF-initiated orthopedic surgery patients as measured through various claim-based measures and PAC assessments. Hospital readmission rates, emergency department visits, and mortality following the anchor hospitalization were not statistically different for patients of PGP-initiated SNF orthopedic surgery episodes than all Model 3 SNF-initiated orthopedic surgery episodes (see Exhibit 118). (Hospital readmission rates and emergency department visits in the 90-day PDP were not produced for this clinical group.)

Exhibit 118: Claims-based Quality Measures for PGP-initiated Orthopedic Surgery Episodes and All SNF Orthopedic Surgery Episodes, Model 3, Q4 2013-Q3 2014

Measure	PGP (N=1,018)			All Model 3 SNF episodes (N=2,273)	Model 3 PGP Risk-adjusted Mean, vs. All Model 3 SNF, Observed Mean
	Risk-adjusted mean	LCI	UCI	Observed Mean	
Unplanned readmission rate, first 30 days of episode	7.0%	5.2%	8.7%	7.2%	-0.2
ED use, first 30 days of episode	4.5%	3.1%	5.8%	4.8%	-0.4
All-cause mortality rate, first 30 days of episode	1.1%	0.3%	1.9%	0.8%	0.3

* Denotes statistical significance at the 5% level.

LCI = lower confidence interval at the 5% level; UCI = upper confidence interval at the 5% level.

Source: Lewin analysis of Medicare claims and enrollment data for episodes that began Q4 2013 through Q3 2014 for BPCI PGP and hospital providers.

E. Impact of BPCI on SNF Non-Surgical Respiratory

1. Participants

During the first year of the initiative, 50 SNFs, or 79% of SNF EIs in Model 3, participated in at least one non-surgical respiratory episode.[71] There were 988 non-surgical respiratory episodes initiated in Model 3-participating SNFs, accounting for 13% of all Model 3 SNF episodes. Approximately 45% of these episodes were for simple pneumonia and respiratory infection MS-DRGs. Each of the SNFs participating in BPCI with these clinical episodes was a non-profit urban facility. They were less likely to be part of a chain than non-participating SNFs (10% vs. 23%). Participating SNFs were also more likely than the typical SNF to have an inpatient rehabilitation facility in their core based statistical area (64% vs. 29%), suggesting more post-acute care options are available in their markets than the markets where there are no participating SNFs.

2. Patient population characteristics

Beneficiaries who were treated in a SNF participating in Model 3 for a non-surgical respiratory episode were similar to all Medicare beneficiaries treated in the same MS-DRGs. Both groups exhibit similar age, gender, and Medicare eligibility due to disability characteristics (see Exhibit 119). The major difference between the BPCI and non-BPCI groups was related to dual eligibility status. A smaller proportion of the patients of BPCI providers were dually eligible compared to all beneficiaries with the same MS-DRG (19.8% vs. 33.6%).

---

[71] The clinical episodes that compose the non-surgical respiratory group are: chronic obstructive pulmonary disease, bronchitis, and asthma; simple pneumonia and respiratory infections; and other respiratory.

Exhibit 119: Characteristics of BPCI Beneficiaries and All Medicare Beneficiaries
with a SNF Admission for a Non-Surgical Respiratory MS-DRG, Model 3, Q4 2013-Q3 2014

Characteristics	Model 3 BPCI SNF Beneficiaries with Non-Surgical Respiratory Episodes, Q4 2013 – Q3 2014 (N = 915)		All Medicare Beneficiaries With SNF Admission following inpatient admission with a Non-surgical Respiratory MS-DRG, Q3 2014 (N = 53,909)	
	N	%	N	%
Age				
20-64	84	9.2%	5,376	10.0%
65-79	333	36.4%	20,007	37.1%
80+	498	54.4%	28,526	52.9%
Gender				
Female	545	59.6%	31,044	57.6%
Male	370	40.4%	22,865	42.4%
Medicaid and Disability				
% Eligible Medicaid	181	19.8%	18,139	33.6%
% Disability, no ESRD	105	11.5%	6,017	11.2%

Source: Lewin analysis of Medicare claims and enrollment data for SNF-initiated episodes that began Q4 2011 through Q3 2014 for BPCI providers and all admissions to non-participating SNFs that had a preceding inpatient stay for the same MS-DRG in Q3 2014.

3. Change in Medicare standardized allowed payments

There were no statistically significant differences in Medicare standardized allowed payments between BPCI episodes and non-BPCI episodes. The average total standardized allowed payment included in the bundle definition for 90 day episodes for patients treated by BPCI providers declined from baseline to intervention periods by 5.8%, but this reduction was not statistically significant relative to the change for patients treated by matched comparison providers (see Exhibit 120). There were also no indications that providers were shifting costs to services outside the bundle period or not covered in the bundle definition.

Exhibit 120: Impact of BPCI on Total Standardized Allowed Payment Included in the Bundle
Definition for Non-Surgical Respiratory Episodes, Model 3 SNFs, Q4 2011-Q3 2014

Measure	Bundle Length	BPCI (N=888)		Comparison (N=885)		Diff-in-Diff estimate		
		Baseline	Intervention	Baseline	Intervention	Diff-in-Diff	LCI	UCI
Total Standardized Allowed Payment Included In The Bundle Definition	90	$32,150	$30,391	$29,150	$28,312	-$920	-$3,340	$1,500

* Denotes statistical significance at the 5% level.

LCI = lower confidence interval at the 5% level; UCI = upper confidence interval at the 5% level.

Note: Baseline is defined as episodes that began Q4 2011 through Q3 2012. Intervention is defined as episodes that began Q4 2013 through Q3 2014. Sample sizes reflect number of episodes initiated during the intervention period that met inclusion criteria for the outcome.

Source: Lewin analysis of Medicare claims and enrollment data for episodes that began Q4 2011 through Q3 2014 for BPCI and comparison providers.

4. Change in beneficiary quality of care

There was no difference in the change in quality of care between BPCI patients with SNF non-surgical respiratory episodes and patients of non-BPCI providers, as measured through various claim-based measures and PAC assessments. The change in hospital readmission rates, emergency department visits, and mortality following the start of the episode for SNF non-surgical respiratory episodes was not statistically different for patients of BPCI providers and comparison providers (see Exhibit 121). There were no statistically significant changes in assessment-based quality measures (i.e. overall function, self-care function and mobility) from baseline to intervention period between BPCI and comparison group beneficiaries with non-surgical respiratory episodes who were treated at a SNF.

Exhibit 121: Impact of BPCI on Claims-based Quality Measures for Non-Surgical Respiratory Episodes, Model 3 SNFs, Q4 2011-Q3 2014

Measure	BPCI (N=905)		Comparison (N=909)		Diff-in-Diff estimate		
	Baseline	Intervention	Baseline	Intervention	Diff-in-Diff	LCI	UCI
Unplanned Readmission Rate, within 30 days of episode start	26.2%	26.0%	24.1%	23.7%	0.3	-4.5	5.0
Emergency Department Visit Rate, within 30 days of episode start	9.2%	9.5%	6.8%	7.6%	-0.5	-4.2	3.2
Mortality, within 30 days of episode start	12.1%	11.1%	9.2%	12.0%	-3.8	-7.5	0.0

* Denotes statistical significance at the 5% level.

LCI = lower confidence interval at the 5% level; UCI = upper confidence interval at the 5% level.

Note: Baseline is defined as episodes that began Q4 2011 through Q3 2012. Intervention is defined as episodes that began Q4 2013 through Q3 2014. Sample sizes reflect number of episodes initiated during the intervention period that met inclusion criteria for the outcome.

Source: Lewin analysis of Medicare claims and enrollment data for episodes that began Q4 2011 through Q3 2014 for BPCI and comparison providers.

F. Impact of BPCI on SNF Non-Surgical Cardiovascular

1. Participants

During the first four quarters of the initiative, 61 skilled nursing facilities (SNF), or 68% of SNF EIs in Model 3, participated in at least one non-surgical cardiovascular clinical episode. There were 1,087 non-surgical cardiovascular episodes initiated in Model 3-participating SNFs, accounting for 15% of all Model 3 SNF episodes.[72] More than half (52%) of these episodes were congestive heart failure MS-DRGs. Each of the SNFs participating in BPCI with these clinical episodes was an urban facility and 85% were non-profits. They were less likely to be of part of a chain than non-participating SNFs (16% vs. 23%). Participating SNFs were also more likely than the typical SNF to have an inpatient rehabilitation facility in their core based statistical area (56% vs. 29%), suggesting

---

[72] The clinical episodes that compose the non-surgical cardiovascular group are: acute myocardial infarction; atherosclerosis; cardiac arrhythmia; chest pain; congestive heart failure; medical peripheral vascular disorders; and syncope & collapse.

more post-acute care options are available in their markets than in markets with no BPCI-participating SNFs.

2. Patient population characteristics

Beneficiaries who were treated in a SNF participating in Model 3 for a non-surgical cardiovascular episode were similar to all Medicare beneficiaries treated in the same MS-DRG. Both groups exhibit similar age, gender, and Medicare eligibility due to disability characteristics (see Exhibit 122). The major difference between the BPCI and non-BPCI groups was related to dual eligibility status. A smaller proportion of the patients of BPCI providers were dually eligible compared to all beneficiaries with the same MS-DRG (18.5% vs. 25.2%).

Exhibit 122: Characteristics of BPCI Beneficiaries and All Medicare Beneficiaries with a SNF Admission for a Non-Surgical Cardiovascular MS-DRG, Model 3, Q4 2013 – Q3 2014

Characteristics	Model 3 BPCI SNF Beneficiaries with Non-Surgical Cardiovascular Episodes, Q4 2013 – Q3 2014 (N = 1,047)		All Medicare Beneficiaries With SNF Admission following inpatient admission with a Non-surgical Cardiovascular MS-DRG, Q3 2014 (N = 47,972)	
	N	%	N	%
Age				
20-64	63	6.0%	2,623	5.5%
65-79	307	29.3%	13,787	28.7%
80+	677	64.7%	31,562	65.8%
Gender				
Female	626	59.8%	29,460	61.4%
Male	421	40.2%	18,512	38.6%
Medicaid and Disability				
% Eligible Medicaid	194	18.5%	12,104	25.2%
% Disability, no ESRD	81	7.7%	2,890	6.0%

Source: Lewin analysis of Medicare claims and enrollment data for SNF-initiated episodes that began Q4 2011 through Q3 2014 for BPCI providers and all admissions to non-participating SNFs that had a preceding inpatient stay for the same MS-DRG in Q3 2014.

3. Change in Medicare standardized allowed payments

SNF EIs were able to decrease the SNF payments during the 90 days post qualifying hospital discharge. The SNF allowed amount decreased from $20,601 to $18,879 for baseline to intervention for BPCI patients. This is a decline of $2,708 more than the difference observed for the comparison group, where the SNF payments increased for the comparison group from $16,305 to $17,292. However, the decrease in SNF payments did not result in a statistically significant decline in the total Medicare standardized allowed payments included in the bundle (see Exhibit 123). The decrease in SNF allowed payment was driven by a statistically significant decline in the number of SNF days. The number of SNF days decreased from 37.7 to 35.1 for baseline to intervention for BPCI patients compared to an increase for the comparison group from 32.5 to 34.3, resulting in a relative

decrease of 4.5 days. There were no indications that these results were achieved by shifting services outside of the bundle definition or bundle period.

Exhibit 123: Impact of BPCI on Total Standardized Allowed Payment Included in the Bundle Definition for Non-Surgical Cardiovascular Episodes, Model 3 SNFs, Q4 2011-Q3 2014

Measure	Bundle Length	BPCI (N=1,012)		Comparison (N=1,001)		Diff-in-Diff estimate		
		Baseline	Intervention	Baseline	Intervention	Diff-in-Diff	LCI	UCI
Total Standardized Allowed Payment Included In The Bundle Definition	90	$33,121	$31,691	$29,465	$28,883	-$848	-$2,872	$1,176

* Denotes statistical significance at the 5% level.

LCI = lower confidence interval at the 5% level; UCI = upper confidence interval at the 5% level.

Note: Baseline is defined as episodes that began Q4 2011 through Q3 2012. Intervention is defined as episodes that began Q4 2013 through Q3 2014. Sample sizes reflect number of episodes initiated during the intervention period that met inclusion criteria for the outcome.

Source: Lewin analysis of Medicare claims and enrollment data for episodes that began Q4 2011 through Q3 2014 for BPCI and comparison providers.

4. *Change in beneficiary quality of care*

BPCI patients with SNF non-surgical cardiovascular episodes had a statistically significant increase in readmission rates from baseline to intervention relative to the patients treated by the matched comparison group (see Exhibit 124). Readmission rates increased from 23.6% to 25.1% for BPCI patients relative to a decrease among patients treated by comparison providers (27.2% to 21.7%) resulting in a relative increase of 7pp. This large increase is due to a high baseline rate for the comparison group that fell by the intervention period and a lower baseline rate for the BPCI providers, which increased by the intervention period. We will continue to monitor this quality measure, although the positive results could be due to anomalies that might not persist. The change in emergency department visits and mortality following the start of the episode for SNF non-surgical cardiovascular episodes were not statistically different for patients of BPCI providers and comparison providers. There were no statistically significant changes in assessment-based quality measures (i.e. overall function, self-care function and mobility) from baseline to intervention period between BPCI and comparison group beneficiaries with non-surgical cardiovascular episodes who were treated at a SNF.

Exhibit 124: Claims-based Quality Measures for Non-Surgical Cardiovascular Episodes, Model 3 SNFs, Q4 2011-Q3 2014

Measure	BPCI (N=1,034)		Comparison (N=1,041)		Diff-in-Diff estimate		
	Baseline	Intervention	Baseline	Intervention	Diff-in-Diff	LCI	UCI
Unplanned Readmission Rate, within 30 days of episode start	23.6%	25.1%	27.2%	21.7%	7.0*	1.9	12.1
Emergency Department Visit Rate, within 30 days of episode start	10.7%	8.6%	9.5%	7.1%	0.3	-2.6	3.3
Mortality, within 30 days of episode start	8.7%	9.9%	9.9%	9.3%	1.8	-1.7	5.2

* Denotes statistical significance at the 5% level.

LCI = lower confidence interval at the 5% level; UCI = upper confidence interval at the 5% level.

Note: Baseline is defined as episodes that began Q4 2011 through Q3 2012. Intervention is defined as episodes that began Q4 2013 through Q3 2014. Sample sizes reflect number of episodes initiated during the intervention period that met inclusion criteria for the outcome.

Source: Lewin analysis of Medicare claims and enrollment data for episodes that began Q4 2011 through Q3 2014 for BPCI and comparison providers.

G. Impact of BPCI on HHA Non-Surgical Respiratory

1. Participants

During the first four quarters of the initiative, 22 HHAs, or 79% of HHA EIs in Model 3, participated in at least one non-surgical respiratory episode. There were 845 non-surgical respiratory episodes initiated in Model 3-participating HHAs, accounting for 26% of all Model 3 episodes. Approximately 44% of these episodes were for chronic obstructive pulmonary disease, bronchitis, or asthma MS-DRGs. Each of the HHAs participating in BPCI with these clinical episodes was a non-profit and 77% were urban facilities. They were smaller than the typical HHA (with an average of 17 employed nurses, compared with 85 nurses in non-participating HHAs).

2. Patient population characteristics

Beneficiaries who were treated in an HHA participating in Model 3 for a non-surgical respiratory episode were similar to all Medicare beneficiaries treated in the same MS-DRG. Both groups were similar with respect to age, dual eligibility status, and disability (see Exhibit 125). The major difference between the BPCI and non-BPCI groups was related to gender. The patients of BPCI HHA providers were more often female than all beneficiaries with the same MS-DRG (63.8% vs. 59.6%).

Exhibit 125: Characteristics of BPCI Beneficiaries and All Medicare Beneficiaries
with an HHA Admission for a Non-Surgical Respiratory MS-DRG, Model 3, Q4 2013 – Q3 2014

Characteristics	Model 3 BPCI HHA Beneficiaries with Non-Surgical Respiratory Episodes, Q4 2013 – Q3 2014 (N = 802)		All Medicare Beneficiaries With HHA following inpatient admission with a Non-surgical Respiratory MS-DRG, Q3 2014 (N = 42,235)	
	N	%	N	%
Age				
20-64	119	14.8%	6,182	14.6%
65-79	339	42.3%	18,003	42.6%
80+	344	42.9%	18,050	42.7%
Gender				
Female	512	63.8%	25,157	59.6%
Male	290	36.2%	17,078	40.4%
Medicaid and Disability				
% Eligible Medicaid	223	27.8	12,142	28.7%
% Disability, no ESRD	153	19.1%	7,020	16.6%

Source: Lewin analysis of Medicare claims and enrollment data for HHA-initiated episodes that began Q4 2011 through Q3 2014 for BPCI providers and all admissions to non-participating HHAs that had a preceding inpatient stay for the same MS-DRG in Q3 2014.

3. Change in Medicare standardized allowed payments

There were no statistically significant differences between BPCI episodes and episodes in comparison providers in Medicare standardized allowed payments. The average total standardized allowed payment included in the bundle definition for 90 day episodes for patients treated by BPCI providers declined from baseline to intervention periods, but this reduction was not statistically significant relative to the change for patients treated by comparison providers (see Exhibit 126). Furthermore, there were no statistically significant differences in the change for the allowed amount for any Part A service (i.e. HH, SNF). There were also no statistically significant differences in the intensity of PAC services (i.e. length of stay or number of home health visits) - relative to the episodes from comparison providers.

Exhibit 126: Total Standardized Allowed Payment Included in the Bundle Definition for Non-Surgical Respiratory Episodes, Model 3 HHAs, Q4 2011-Q3 2014

Measure	Bundle Length	BPCI (N=789)		Comparison (N=783)		Diff-in-Diff estimate		
		Baseline	Intervention	Baseline	Intervention	Diff-in-Diff	LCI	UCI
Total Standardized Allowed Payment Included In The Bundle Definition	90	$14,042	$13,062	$13,907	$13,045	-$118	-$2,024	$1,788

* Denotes statistical significance at the 5% level.

LCI = lower confidence interval at the 5% level; UCI = upper confidence interval at the 5% level.

Note: Baseline is defined as episodes that began Q4 2011 through Q3 2012. Intervention is defined as episodes that began Q4 2013 through Q3 2014. Sample sizes reflect number of episodes initiated during the intervention period that met inclusion criteria for the outcome.

Source: Lewin analysis of Medicare claims and enrollment data for episodes that began Q4 2011 through Q3 2014 for BPCI and comparison providers.

4. *Change in beneficiary quality of care*

The quality of care for BPCI patients with HHA non-surgical respiratory episodes was similar to that of patients of non-BPCI providers, as measured through various claim-based measures. The change in hospital readmission rates, emergency department visits, and mortality following the start of the episode for HHA non-surgical respiratory episodes were not statistically different for patients of BPCI providers and comparison providers (see Exhibit 127). There were no statistically significant changes in assessment-based quality measures (i.e. bathing, ambulation, upper/lower body dressing and bed transferring) from baseline to intervention period between BPCI and comparison group beneficiaries with non-surgical respiratory episodes who were treated at a HHA.

Exhibit 127: Claims-based Quality Measures for Non-Surgical Respiratory Episodes, Model 3 HHAs, Q4 2011-Q3 2014

Measure	BPCI (N=796)		Comparison (N=792)		Diff-in-Diff estimate		
	Baseline	Intervention	Baseline	Intervention	Diff-in-Diff	LCI	UCI
Unplanned Readmission Rate, within 30 days of episode start	20.2%	19.0%	17.8%	18.6%	-2.1	-6.1	2.0
Emergency Department Visit Rate, within 30 days of episode start	15.1%	12.1%	8.6%	7.3%	-1.6	-5.2	1.9
Mortality, within 30 days of episode start	4.6%	3.9%	3.5%	4.9%	-2.2	-4.7	0.3

* Denotes statistical significance at the 5% level.

LCI = lower confidence interval at the 5% level; UCI = upper confidence interval at the 5% level.

Note: Baseline is defined as episodes that began Q4 2011 through Q3 2012. Intervention is defined as episodes that began Q4 2013 through Q3 2014. Sample sizes reflect number of episodes initiated during the intervention period that met inclusion criteria for the outcome.

Source: Lewin analysis of Medicare claims and enrollment data for episodes that began Q4 2011 through Q3 2014 for BPCI and comparison providers.

H. Impact of BPCI on HHA Non-Surgical Cardiovascular

1. *Participants*

During the first four quarters of the initiative, 27 HHAs, or 96% of HHA EIs in Model 3, participated in at least one non-surgical cardiovascular episode. There were 1,485 non-surgical cardiovascular episodes initiated in Model 3-participating HHAs, accounting for 46% of all Model 3 episodes. Approximately 70% of these episodes were for congestive heart failure MS-DRGs. Many of the HHAs participating in BPCI with these clinical episodes were urban facilities (78%) and 4% were non-profits. They were somewhat smaller than the typical HHA (with an average of 72 employed nurses, compared with 85 employed nurses in non-participating HHAs).

2. *Patient population characteristics*

Beneficiaries who were treated in an HHA participating in Model 3 for a non-surgical cardiovascular episode were similar to all Medicare beneficiaries treated in the same MS-DRG. Both groups were similar with respect to age, gender, and disability. The major difference between the BPCI and non-BPCI groups was related to dual eligibility status. The patients of BPCI providers were more likely to be dual insured than all beneficiaries with the same MS-DRG (31.6% vs. 23.0%; see Exhibit 128).

Exhibit 128: Characteristics of BPCI Beneficiaries and All Medicare Beneficiaries with an HHA Admission for a Non-Surgical Cardiovascular MS-DRG, Model 3, Q4 2013 – Q3 2014

Characteristics	Model 3 BPCI HHA Beneficiaries with Non-Surgical Cardiovascular Episodes, Q4 2013 – Q3 2014 (N = 1,485)		All Medicare Beneficiaries With HHA following inpatient admission with a Non-surgical Cardiovascular MS-DRG, Q3 2014 (N = 44,762)	
	N	%	N	%
Age				
20-64	101	7.0%	3,740	8.4%
65-79	482	33.2%	15,077	33.7%
80+	867	59.8%	25,945	58.0%
Gender				
Female	888	61.2%	26,538	59.3%
Male	562	38.8%	18,224	40.7%
Medicaid and Disability				
% Eligible Medicaid	458	31.6%	10,278	23.0%
% Disability, no ESRD	126	8.7%	4,170	9.3%

Source: Lewin analysis of Medicare claims and enrollment data for HHA-initiated episodes that began Q4 2011 through Q3 2014 for BPCI providers and all admissions to non-participating HHAs that had a preceding inpatient stay for the same MS-DRG in Q3 2014.

3. Change in Medicare standardized allowed payment amounts

There were no statistically significant differences between BPCI episodes and episodes in matched providers in the change from the baseline to the intervention periods in Medicare standardized allowed payments. The average total standardized allowed payment included in the bundle definition for 90 day episodes for patients treated by BPCI providers declined from baseline to intervention periods, but this reduction was not statistically significant relative to the change for patients treated by matched comparison providers (see Exhibit 129). Furthermore, there were no statistically significant differences in the change for the allowed amount for any Part A service (i.e. HH, SNF). This is despite the fact there was a statistically significant decrease in the number of HHA visits (2.3 visits) and a statistically significant decrease in the number of SNF days (3 days) relative to the matched comparison group. There were no indications that these results were achieved by shifting services outside of the bundle definition or bundle period.

Exhibit 129: Impact of BPCI on Total Standardized Allowed Payment Included in the Bundle Definition for Non-Surgical Cardiovascular Episodes, Model 3 HHAs, Q4 2011-Q3 2014

Measure	Bundle Length	BPCI (N=1,394)		Comparison (N=1,427)		Diff-in-Diff estimate		
		Baseline	Intervention	Baseline	Intervention	Diff-in-Diff	LCI	UCI
Total Standardized Allowed Payment Included In The Bundle Definition	90	$15,531	$15,361	$14,610	$14,637	-$196	-$1,320	$928

* Denotes statistical significance at the 5% level.

LCI = lower confidence interval at the 5% level; UCI = upper confidence interval at the 5% level.

Note: Baseline is defined as episodes that began Q4 2011 through Q3 2012. Intervention is defined as episodes that began Q4 2013 through Q3 2014. Sample sizes reflect number of episodes initiated during the intervention period that met inclusion criteria for the outcome.

Source: Lewin analysis of Medicare claims and enrollment data for episodes that began Q4 2011 through Q3 2014 for BPCI and comparison providers.

4. *Change in beneficiary quality of care*

BPCI patients with HHA non-surgical cardiovascular episodes had an increase in readmission rates from baseline to intervention relative to the patients treated by the matched comparison group. Readmission rates increased from 23.7% to 25.1% for BPCI patients relative to a decrease among patients treated by comparison providers (24.3% to 21.4%), resulting in a statistically significant relative increase of 4.2pp (see Exhibit 130). The change in emergency department visits and mortality following the start of the episode for HHA non-surgical cardiovascular episodes were not statistically different for patients of BPCI providers and comparison providers. There were no statistically significant changes in assessment-based quality measures (i.e. bathing, ambulation, upper/lower body dressing and bed transferring) from baseline to intervention period between BPCI and comparison group beneficiaries with non-surgical cardiovascular episodes who were treated at a HHA.

Exhibit 130: Impact of BPCI on Claims-based Quality Measures for Non-Surgical Cardiovascular Episodes, Model 3 HHAs, Q4 2011-Q3 2014

Measure	BPCI (N=1,412)		Comparison (N=1,437)		Diff-in-Diff estimate		
	Baseline	Intervention	Baseline	Intervention	Diff-in-Diff	LCI	UCI
Unplanned Readmission Rate, within 30 days of episode start	23.7%	25.1%	24.3%	21.4%	4.2*	0.9	7.5
Emergency Department Visit Rate, within 30 days of episode start	11.9%	11.3%	8.1%	7.3%	0.3	-2.3	2.8
Mortality, within 30 days of episode start	4.0%	3.7%	3.7%	3.9%	-0.5	-2.5	1.6

* Denotes statistical significance at the 5% level.

LCI = lower confidence interval at the 5% level; UCI = upper confidence interval at the 5% level.

Note: Baseline is defined as episodes that began Q4 2011 through Q3 2012. Intervention is defined as episodes that began Q4 2013 through Q3 2014. Sample sizes reflect number of episodes initiated during the intervention period that met inclusion criteria for the outcome.

Source: Lewin analysis of Medicare claims and enrollment data for episodes that began Q4 2011 through Q3 2014 for BPCI and comparison providers.

I.　Physician Group Practice (PGP) HHA-Initiated Episodes

During the first year of the initiative, there were only 109 Model 3 PGP HHA-initiated orthopedic surgery episodes, which did not support DiD analyses relative to a non-BPCI comparison group of HHAs.[73] However, similar to the PGP SNF initiated episodes, we were able to compare the Model 3 PGP orthopedic surgery episodes initiated in a HHA (n=251) to all Model 3 orthopedic surgery episodes initiated in a HHA (n=360). Given the difficulty linking individual physicians with their PGPs, these results should be viewed as preliminary and with caution.

Under Model 3, HHA beneficiaries treated by PGPs achieved similar quality, utilization, and payment outcomes relative to all Model 3 HHA-initiated orthopedic surgery episodes.

The characteristics of the PGP HHA-initiated orthopedic surgery episode patients were similar to all Model 3 HHA orthopedic surgery episode patients with regards to gender, Medicaid eligibility, disability status, and average MS-DRG case weight. The PGP HHA-initiated orthopedic surgery episode patients were less likely to have received home health care (10.0% vs. 13.3%) or have been hospitalized in an acute care setting (8.0% vs. 10.8%) prior to the qualifying inpatient hospitalization. Additionally, PGP HHA-initiated orthopedic surgery episode patients were less likely to be over 80 years of age (12.7% vs. 15.8%) and had a lower average HCC index than all Model 3 HHA orthopedic surgery episode patients (0.205 vs. 0.308). The differences in prior health care utilization and HCC index were also present during the four quarters prior to the BPCI initiative.

Despite differences in patient characteristics, HHA beneficiaries treated by PGPs achieved similar quality, utilization, and payment outcomes relative to all Model 3 HHA-initiated orthopedic surgery episodes. The total standardized allowed payment included in the bundle definition was no different for PGP-initiated orthopedic surgery episodes compared to all Model 3 HHA orthopedic surgery episodes. Hospital readmission rates, emergency department visits, and mortality within the first 30 days of the episode were not statistically different for PGP beneficiaries relative to all Model 3 HHA BPCI beneficiaries in orthopedic surgery episodes. (Hospital readmission rates and emergency department visits in the 90-day PDP were not produced for this clinical group.)

---

[73] Please note that the data linking individual physicians with their PGP contained errors, so these results should be viewed with caution.

Exhibit 131: Total Standardized Allowed Payment Included in the Bundle Definition for PGP-initiated Orthopedic Surgery Episodes and All HHA Orthopedic Surgery Episodes, Model 3, Q4 2013-Q3 2014

Measure	Bundle Length	PGP (N=235)			All Model 3 HHA episodes (N=252)	Model 3 PGP Risk-adjusted Mean, vs. All Model 3 HHA, Observed Mean
		Risk-adjusted mean	LCI	UCI	Observed Mean	
Total Standardized Allowed Payment Included In The Bundle	90	$6,807	$6,006	$7,608	$6,363	$444

* Denotes statistical significance at the 5% level.

LCI = lower confidence interval at the 5% level; UCI = upper confidence interval at the 5% level.

Source: Lewin analysis of Medicare claims and enrollment data for episodes that began Q4 2013 through Q3 2014 for BPCI PGP and HHA providers.

V. Model 4 Results

This section presents information about the experience of Model 4 BPCI providers and their episodes of care, organized by research question. Similar to Models 2 and 3, quantitative analyses of Medicare claims data are based on episodes initiated during the first full year of BPCI (Q4 2013 through Q3 2014); beneficiary assessment data are based on episodes initiated during the first three quarters of BPCI (Q4 2013 through Q2 2014); Awardee-submitted data are based on calendar year 2014 or Q4 2014 depending on the measure; and qualitative data from interviews reflect the first six quarters of the BPCI initiative (Q4 2013 through Q1 2015) while site visits results reflect the first seven quarters of the BPCI initiative (Q4 2013 through Q2 2015). Model 4 BPCI beneficiaries were not surveyed because there were too few of them to allow for a sufficient sample size for analysis. The quantitative outcomes are risk-adjusted as described in Section II.D.2 above.

A. Characteristics of the Initiative and Participants

1. Participants

This section discusses the characteristics of the Model 4 BPCI-participating hospitals that joined the initiative through Q3 2014. There were 13 Model 4 Awardees with 20 EI hospitals that participated in BPCI at some point during the first year of the initiative. However, it should be noted that 10 of the 20 EIs terminated their participation in BPCI prior to September 2014. See **Appendix Q** for a detailed description of the growth of the initiative through July 2015.

Exhibit 132 compares the first year Model 4 BPCI-participating hospital EIs to non-participating hospitals. In regard to ownership, most BPCI-participating hospitals in Model 4 were non-profits (65%). On average, 50% of Model 4 participating hospitals were part of a chain, which was similar to non-participating hospitals (52%).

All BPCI-participating hospitals were located in urban areas, compared with 71% of non-participating hospitals. The average bed count for participating hospitals was more than double that of non-participants (427 vs. 188). Participating hospitals had higher average resident-to-bed ratio than non-participating hospitals (0.14 vs. 0.06) and had over twice as admissions for MS-DRGs included in BPCI episodes during 2011 (4,516 vs. 2,140). The urban locations, higher bed counts, and higher resident-to-bed ratios, may signify a higher concentration of academic hospitals involved in BPCI than in non-participating hospitals.

Compared with non-participating hospitals, BPCI hospitals had a lower share of Medicare days. Participating hospitals had smaller average proportions of Medicare days as a percent of total inpatient days compared to non-participating hospitals (29% vs. 41%). BPCI-participating hospitals had similar disproportionate share percentages as non-participating hospitals (29% vs. 28%), indicating the same proportion of Medicare Supplemental Security Income, Medicaid, or other low-income populations.

Model 4 BPCI-participating hospital EI cardiovascular surgery episodes had 2% higher standardized payments for the inpatient stay plus the 90 day PDP in 2011 relative to episodes from non-participating hospitals. The standardized allowed payments for the inpatient stay plus the 90 day PDP for orthopedic surgery episodes initiated at Model 4 BPCI-participating hospitals, however, was 5% lower than for episodes at non-participating hospitals.

Exhibit 132: Characteristics of BPCI-participating Hospital EIs and Non-participating Hospitals, Model 4, Q4 2013 - Q3 2014

	BPCI Hospital EIs (N=20)		Non-participating Hospitals (N=3,056)	
	N	%	N	%
Ownership				
For Profit	6	30	672	22
Government	1	5	550	18
Non-Profit	13	65	1,803	59
Urban/Rural				
Rural	0	0	886	29
Urban	20	100	2,170	71
Part of Chain				
Yes	10	50	1,528	52
	Mean		**Mean**	
Bed Count	427		188	
Number of BPCI Episodes	4,516		2,140	
Medicare Days Percent	29%		41%	
Resident-bed ratio	0.14		0.06	
Disproportionate Share Percent	29%		28%	
Standardized Part A Allowed Payment inpatient stay plus 90 day PDP, 2011				
Clinical Episode Group (N, BPCI discharges)	**Mean**		**Mean**	
Orthopedic surgery (5,099)	$27,540		$28,947	
Cardiovascular surgery (7,603)	$32,944		$32,448	

Source: Lewin analysis of 2013 Provider of Service (POS) files and 2011 Medicare claims. BPCI participating hospitals are defined as Model 4 Episode Initiators, Q4 2013 – Q3 2014. Non-Participant hospitals are all other hospitals and exclude Model 2 hospitals participating in BPCI during the first year.

a. *Awardee submitted baseline characteristics*

Awardees and EIs submitted information about their experiences that may have contributed to their implementation of BPCI through an on-line data reporting tool beginning in Q1 2015. This included patient payer mix, prior experience with care redesign and payment initiatives, and quality of care indicators. We collected data representing the Awardee and EI experience during various time periods, including prior to joining BPCI and during 2014.

Patients by payer type

We calculated patient mix for 14 Model 4 EIs. Payer mix was based on the reported number of beneficiaries in 2014 with a primary payer of Medicaid, Medicare, private, or other. As shown in Exhibit 133, Medicare beneficiaries comprised the highest portion of total beneficiaries for Model 4 EIs, followed by private payer beneficiaries. The mean percent of Medicare beneficiaries and private payer beneficiaries were 43.9% and 37.9%, respectively.

Exhibit 133: Patient Payer Mix, Model 4, 2014

	Model 4 EIs N=14			
	Mean	Min	Median	Max
Percentage of Medicaid Patients	13.1	1.0	13.5	28.4
Percentage of Medicare Patients	43.9	24.2	39.0	74.3
Percentage of Private Payer Patients	37.9	6.3	37.5	67.2
Percentage of Other Payer Patients	5.0	0.0	3.5	13.3

Source: Lewin analysis of Awardee-submitted data collected February 2015 for Model 4 episode initiators participating in BPCI between Q4 2013 – Q4 2014.

Prior experience

As illustrated in Exhibit 134, Model 4 EIs reported whether they participated in care redesign or payment incentives initiatives prior to their participation in BPCI. Enhancement of care delivery was the most commonly reported care redesign experience for this model, reported by 85.7% of the participants. The majority of participants also reported prior participation in the redesign of care pathways (78.6%), as well as patient activation, engagement and risk management activities (78.6%).

The most common type of prior payment incentives experience among Model 4 EIs was pay for performance, with 85.7% of participants reporting prior participation. Conversely, a minority reported experience in bundled payments (14.3%).

Exhibit 134: Care Redesign and Payment Incentives Experience, Model 4, Prior to BPCI Participation

	Model 4 EIs N=14	
	N	%
Prior experience in care redesign initiatives:		
Redesign of Care Pathways	11	78.6
Enhancements in Care Delivery	12	85.7
Patient Activation, Engagement & Risk Management	11	78.6
Care Coordination	10	71.4
System Changes to Support Care	10	71.4
Other Redesign Activities	1	7.7
Prior experience in payment incentives:		
Bundled Payments	2	14.3
Pay for Performance	12	85.7
Shared Savings	4	28.6
Other Payment Incentives	4	28.6

Source: Lewin analysis of Awardee-submitted data collected February 2015 for Model 4 episode initiators participating in BPCI between Q4 2013 – Q4 2014.

Electronic health record (EHR) use

All 14 of the Model 4 participants from which we collected data reported use of EHRs. Among these EIs, 100% reported having the following three functionalities: e-Prescribing, computerized physician order entry, and discharge instructions and care summary documents. Twelve of these EIs (85.7%) also reported medication management and clinical decision support as functionalities of their EHR systems.

Exhibit 135: Electronic Health Record (EHR) Use, Model 4, 2014

	Model 4 EIs N=14	
	N	%
EIs with an EHR	14	100.0
Meaningful-use functionalities (among EIs with EHRs):		
Automated Quality Reporting	10	71.4
Discharge Instructions and Care Summary Documents	14	100.0
Medication Management	12	85.7
e-Prescribing	14	100.0
Computerized Physician Order Entry	14	100.0
Clinical Decision Support	12	85.7

Source: Lewin analysis of Awardee-submitted data collected February 2015 for Model 4 episode initiators participating in BPCI between Q4 2013 – Q4 2014.

EIs that used EHRs were also asked to report the health information exchange (HIE) capabilities of their EHR systems. Of the 14 Model 4 EIs with EHRs, 100% had HIE capabilities. All of the EIs reported using these capabilities to exchange information with other health care providers. The majority (92.9%) also reported using the HIE capabilities to provide information to patients.

Exhibit 136: Health Information Exchange (HIE) Capabilities among Episode Initiators with EHRs, Model 4, 2014

	Model 4 EIs N=14	
	N	%
EIs using EHRs that have HIE Capabilities	14	100.0%
Use of HIE capabilities to exchange data with:		
Providers	14	100.0%
Patients	13	92.9%
None of the Above	0	0.0%

Source: Lewin analysis of Awardee-submitted data collected February 2015 for Model 4 episode initiators participating in BPCI between Q4 2013 – Q4 2014.

b.   *Awardee submitted quality measures*

Medication reconciliation at discharge

Model 4 EIs reported the number of BPCI-beneficiary discharges among beneficiaries 65 years of age and older for which medications were reconciled at discharge and the total number of discharges among this population during Q4 2014. As displayed in Exhibit 137, across EIs, the mean proportion of BPCI beneficiaries that received medication reconciliation at discharge was 84%, and the median percentage was 100%.

Exhibit 137: Proportion of BPCI Beneficiaries Age 65+ that received Medication Reconciliation at Discharge, Model 4, Q4 2014

	Model 4 EIs (N=10)			
	Mean	**Median**	**Minimum**	**Maximum**
Proportion of BPCI Beneficiaries age 65+ that received medication reconciliation at discharge	84%	100%	7%	100%

Source: Lewin analysis of Awardee-submitted data collected February 2015 for Model 4 episode initiators participating in BPCI during Q4 2014.

FDA reportable events

Model 4 EIs report any BPCI beneficiary deaths or serious injuries that are reportable to the FDA that occurred during Q4 2014. As displayed in Exhibit 138, no Model 4 EIs reported any adverse events among BPCI beneficiaries.

Exhibit 138: Proportion of BPCI Beneficiaries that Experienced a Patient Death or Serious Injuries Reportable to the FDA, Model 4, Q4 2014

	Patient Death or Serious Injuries Reportable to the FDA by Model 4 EIs (N=14)			
	Mean	**Median**	**Minimum**	**Maximum**
Proportion of BPCI Beneficiaries that experienced a Patient death/injury associated with the use of the contaminated drugs, devices, or biologics provided by the health care setting	0%	0%	0%	0%
Proportion of BPCI Beneficiaries with an injury due to the use/function of a device in patient care, in which the device is used or functions other than as intended	0%	0%	0%	0%

Source: Lewin analysis of Awardee-submitted data collected February 2015 for Model 4 episode initiators participating in BPCI during Q4 2014.

2.   *Market characteristics*

The Model 2 result section above includes a comparison of the BPCI markets to non-BPCI markets based on the Model 2 and 4 BPCI participants. See Section III.A.2 for a summary of the results.

3.   Model incentive structure characteristics

a.   Entry decisions

The Awardee interviews and case studies provided insights into the reasons that organizations chose to participate in Model 4 of BPCI. We conducted 3 case studies and 10 Awardee interviews with Model 4 participants. Generally speaking, entry decisions were made primarily by the hospitals' administrative leadership, with guidance from their physician leaders.

Model 4 Awardees that participated in quarterly interviews said that they joined the initiative because they saw potential financial opportunities and had identified areas to lower costs, such as reducing length of stay or implant costs. Two Awardees said that the ability to align incentives through gainsharing made the initiative attractive. One of these Awardees elaborated further, noting that gainsharing is particularly appealing because the Awardee's state essentially prohibits hospitals from employing physicians.

Two Model 4 EIs interviewed during case studies this past year indicated that they joined BPCI in order to learn more about bundled payments, which they expect to be a requirement of future payment reforms. In discussing why they chose Model 4 over Model 2, both cited predictability

> *"Model 2 has too many risks without the control of the PAC."*

as a key factor, which is consistent with findings from the first year of the evaluation. Participants felt they had more control over the risks under this model. A common concern heard from Model 4 participants was that Model 2 required strong PAC relationships. One interviewee stated that they would prefer to stick to their areas of strength rather than trying to develop these relationships.

Quality considerations continued to be important for Model 4 participants. Two Awardees said they joined the initiative because they saw BPCI as an opportunity to improve the quality of care, and added that they could do so in a "cost-effective manner." During a case study, another EI stated that they focused on quality metrics that they could quickly impact to see immediate returns.

> *"50% of the change in Model 4 is changing the payment system and 50% is care redesign."*

Participants noted that physician engagement was an important consideration in their entry decision. For one Awardee, they felt Model 4 was the only model that would allow them to receive physician buy-in. Physician engagement was particularly important among Model 4 interviewees, who consulted physicians when selecting episodes. Case studies with Model 4 hospitals also confirmed the importance of provider engagement. Multiple participants said a key factor in successful implementation of the BPCI initiative was the presence of a physician champion, consistent with what was heard in the first year of the evaluation.

b.   Episode and length selection

Clinical episode selection was much narrower for Model 4 participants than those in Model 2 or 3 of BPCI. During the first year of the initiative, Model 4 EIs participated in 17 out of 48 clinical episodes, representing four out of nine clinical episode groups. Seventy percent of EIs participated in major joint replacement of the lower extremity, making it the most common clinical episode among Model 4 participants. Coronary artery bypass graft and double joint replacement of the lower extremity were the next most common clinical episodes, each chosen by 45% of EIs during the first year of the initiative. EI participation was less than 10% in 33 clinical

episodes. The count of EIs participating in each of the 48 clinical episodes during the first year of BPCI is shown in Exhibit 139.

Exhibit 139: Participation of Episode Initiators by Clinical Episode, Model 4, Q4 2013 - Q3 2014

Clinical Episode	Model 4 Episode Initiators (N=20)	
	N	**%**
Non-surgical and Surgical: Gastrointestinal (GI)		
Esophagitis, gastroenteritis and other digestive disorders	0	0
Gastrointestinal hemorrhage	0	0
Gastrointestinal obstruction	0	0
Major bowel procedure	0	0
Total	**0**	**0**
Non-surgical: Cardiovascular		
Acute myocardial infarction	0	0
Atherosclerosis	0	0
Cardiac arrhythmia	0	0
Chest pain	0	0
Congestive heart failure	1	5
Medical peripheral vascular disorders	0	0
Syncope & collapse	0	0
Total	**1**	**5**
Non-surgical Neurovascular		
Stroke	0	0
Transient ischemia	0	0
Total	**0**	**0**
Non-surgical Orthopedic		
Fractures of the femur and hip or pelvis	0	0
Medical non-infectious orthopedic	0	0
Total	**0**	**0**
Non-surgical: Other Medical		
Cellulitis	0	0
Diabetes	0	0
Nutritional and metabolic disorders	0	0
Red blood cell disorders	0	0
Renal failure	0	0
Sepsis	0	0
Urinary tract infection	0	0
Total	**0**	**0**

Clinical Episode	Model 4 Episode Initiators (N=20)	
	N	**%**
Non-surgical: Respiratory		
Chronic obstructive pulmonary disease, bronchitis, asthma	*0*	*0*
Other respiratory	0	0
Simple pneumonia and respiratory infections	0	0
Total	**0**	**0**
Cardiovascular surgery		
AICD generator or lead	1	5
Cardiac defibrillator	7	35
Cardiac valve	6	30
Coronary artery bypass graft	9	45
Major cardiovascular procedure	0	0
Other vascular surgery	0	0
Pacemaker	7	35
Pacemaker device replacement or revision	6	30
Percutaneous coronary intervention	7	35
Total	**10**	**50**
Orthopedic surgery		
Amputation	0	0
Double joint replacement of the lower extremity	9	45
Hip & femur procedures except major joint	0	0
Lower extremity and humerus procedure except hip, foot, femur	0	0
Major joint replacement of the lower extremity	14	70
Major joint replacement of the upper extremity	0	0
Other knee procedures	2	10
Removal of orthopedic devices	0	0
Revision of the hip or knee	3	15
Total	**14**	**70**
Spinal surgery		
Back & neck except spinal fusion	4	20
Cervical spinal fusion	4	20
Combined anterior posterior spinal fusion	2	10
Complex non-cervical spinal fusion	2	10
Spinal fusion (non-cervical)	4	20
Total	**4**	**20**

Source: Lewin Analysis of CMS' BPCI database, June 2015.

Note: The total number of episode initiators for a given clinical group will not add up to the count of episode initiators participating in each of the clinical episodes within that group because episode initiators can participate in more than one clinical episode within the group.

In interviews, participants shared their rationales for selecting episodes. Awardees often selected episodes based on the opportunity to improve quality and reduce costs. In order to identify these opportunities, two Awardees consulted with outside partners who conducted data analysis, showing which episodes seemed to be "low hanging fruit" where savings could be realized. Another factor noted by the majority of Model 4 participants was the presence of highly engaged physician champions who supported the implementation of BPCI for certain episodes. EIs at both case studies conducted in the past year specifically mentioned that their episode selection was influenced by the presence of physicians who were willing and eager to participate in BPCI. Related to this, participants also selected episodes that were part of key service lines in their organization where they had already developed the skill sets needed to be successful.

Model 4 participants cited challenges with respect to the way BPCI episodes are defined. Some indicated that episode definitions are too broad and could include beneficiaries who did not fit into the new care pathways designed for that bundle. For instance, one participant noted that the care redesign they implemented for joint replacement beneficiaries did not necessarily apply to beneficiaries undergoing non-elective surgery because it included interventions such as pre-surgical education classes. During an interview, one Model 4 participant that terminated participation in the initiative cited the inclusion of non-elective joint replacement cases as a factor in its decision to withdraw; another withdrawn participant expressed the opinion that heart transplants should not be included in the CHF bundle due to their rarity. Several participants also noted challenges related to short stay exclusions, explaining that when the inpatient length of stay is shorter than average, beneficiaries may be excluded from the bundle.

c. Partners

Model 4 participants continue to partner informally (e.g., without a contract or formal agreement) with PAC providers to improve patient care under the BPCI program. Model 4 participants indicated that partnering with PAC providers supported communication and coordination from the ACH to the PAC. This communication, however, varied by participant. One Awardee described their SNF partner as being receptive to more frequent and open communication. Another participant described a proactive strategy to reach out to local PAC facilities that are most commonly used to better understand how the PAC providers operate and better understand that element of patient care. Model 4 participants also described partnerships with contractors to conduct data analytics, administer the BPCI program, and distribute any gains through gainsharing. In addition, some Model 4 participants have partnered with physician groups or independent physicians to assist them with claims issues that may arise through the BPCI program. One Awardee has its administrative staff reach out to independent physician practices to ensure they are appropriately paid for the services they provide. Further, Model 4 Awardees have experienced higher levels of physician engagement as a result of BPCI.

> *"I think the partners play a key role. We depend on them. I think we couldn't do it without them."*

Model 4 participants described several benefits to partnership, including improved communication across the care continuum, coordination of efforts and accountability under BPCI, and alignment of physician incentives. One Awardee specifically described how partnerships have improved care redesign. This Awardee partnered with various groups involved in their total joint replacement episode, including the orthopedic surgeons and anesthesia group. This multidisciplinary team positively impacted beneficiary outcomes through concerted efforts at

> *"If you didn't have a [third party administrator], this would be very, very arduous, very difficult to monitor for a hospital because we're not used to being the payer."*

every phase of care. The Awardee described its success in terms of enhanced pain management, reduced length of stay, and identifying best practices for pain management. Other Awardees affirmed that the success of the initiative would not be possible without their partners. One Awardee stated that their informal partnerships are just as important as their contractual relationships; this Awardee has successfully collaborated with SNFs in its area to reinforce standards of care and prevent readmissions.

There were also challenges associated with partnerships. Participants explained that partnering with physicians to standardize care has been difficult. Despite gainsharing and providing data on patient outcomes, it is challenging to partner with and standardize practices among physicians who have varying approaches to care provision. Another challenge described by participants is the inability to partner with every PAC provider to which their patients are discharged. Despite efforts to build relationships with regional PAC providers, there are instances where beneficiaries cannot be accounted for during the post-discharge period; these beneficiaries present a higher likelihood for readmission.

d.   *Waiver use*

Based on a review of Awardee Implementation Protocols, approximately 94% of EIs in Q3 2014 intended to participate in gainsharing, while none of the EIs were offering beneficiary incentives (See Exhibit 140).[74] No other waivers were available to Model 4 participants. For the Q3 2014 Awardees, if the Awardee was participating in a waiver, all EIs participated as well.

Exhibit 140: Participation of Episode Initiators in BPCI Waivers, Model 4, Q3 2014

	Model 4 EIs (N=16)			
	Model 4 EIs Allowed to Use Waiver		**Model 4 EIs that Used Waiver**	
Model 4 Waivers	**N**	**%**	**N**	**%**
Beneficiary Incentives	0	0.0%	0	0.0%
Gainsharing	15	93.8%	NA*	NA*

*No data are available regarding use of the gainsharing waiver.

Source: Lewin analysis of Awardee Implementation Protocols for the Awardee active in the BPCI initiative in Q3 2014.

Note: The 16 EIs in Q3 2014 are distributed among 12 Awardees. Due to a delay in collecting information from Awardees regarding gainsharing, we cannot comment on how often this waiver has been used during the first year of the initiative.

Gainsharing

The gainsharing waiver allows BPCI participants to share incentive payments, with limitations, with partners that have entered into a gainsharing agreement. In the fourth quarter of the

---

[74] For a description of the waivers, see Section A.2 BPCI Initiative.

initiative, 15 of the 16 Model 4 providers elected the gainsharing waiver. Of these facilities, 11 intend to share gains with physicians.[75]

BPCI providers continue to see gainsharing as a way to foster engagement, particularly among physicians. During a case study, one site noted that BPCI supplemented the quality improvement program they already had in place by offering financial incentives to physicians. The site explained that physician buy-in was essential to success in BPCI because physicians know where efficiencies can be realized. This participant involved physicians from the beginning of the initiative, allowing them to select measures that would be used to determine gainsharing eligibility. Though gains had not been disbursed at the time of the case study, the site indicated that several surgeons and cardiologists are eligible for gains.

Additionally, gainsharing can encourage communication and coordination among physicians. During an interview, one Awardee noted that they adopted a "holistic" approach to gainsharing by determining eligibility for gains on a case-basis rather than a physician-basis. The Awardee reported that physicians and surgeons of various disciplines have started to "treat the case as a whole," leading to more cohesive and comprehensive care. This site reported that the gainsharing program is very well received by their physician representatives.

Other sites revealed that their physicians were less receptive to the gainsharing waiver. One EI that participated in a case study noted that their physicians were initially reluctant to enter into gainsharing agreements with the hospital, citing ambiguity over which party would cover administrative costs. Ultimately, after further education and encouragement, nearly 100 physicians accepted the terms of the gainsharing agreement. The Awardee noted that since these agreements have been established, physicians in other practice groups have expressed a desire to participate in gainsharing. The hospital noted that they are considering additional BPCI episodes to include these physicians in the initiative.

4. Care redesign and cost saving strategy characteristics

During the first Awardee data submission period, Model 4 EIs reported if they participated in five types of care redesign activities and the status of each intervention during 2014. As displayed in Exhibit 141, the majority of EIs were participating in all five types of care redesign. Participation was higher in interventions related to redesign of care pathways (93%), enhancements in care delivery (93%), and care coordination (93%) than participation in activities related to patient activation, engagement, and risk management (86%) and system changes to support care (86%). Out of the EIs who participated in a care redesign activity, very few reported their care redesign activity as completed and terminated. None of the participating EIs had completed and terminated redesign of care pathways, enhancement in care delivery, and system changes to support care, while only eight percent completed and terminated patient activation, engagement, and risk management and care coordination activities.

---

[75] Due to a delay in collecting information from Awardees regarding gainsharing, we cannot comment on how often this waiver has been used during the first year of the initiative.

Exhibit 141: Participation and Status of Care Redesign Interventions, Model 4, 2014

	Redesign of Care Pathways	Enhancement in Care Delivery	Patient Activation, Engagement, and Risk Management	Care Coordination	System Changes to Support Care
Percentage of EIs that indicated participation	93%	93%	86%	93%	86%
Care redesign status (among participating EIs)					
In planning stage	46%	31%	17%	31%	42%
Implementation started	31%	31%	42%	31%	33%
Fully operational	23%	38%	33%	31%	25%
Completed & terminated	0%	0%	8%	8%	0%

Source: Lewin analysis of Awardee-submitted data collected February 2015 for 2014 Model 4 BPCI participants.

a. Care redesign and care coordination

The case studies and quarterly calls provided more details on how participants were implementing care redesign activities across the five categories described above. Many described activities related to patient education, standardization of care, and increased care coordination using case managers/care navigators, all with the goal of reducing readmissions to the hospital.

Several Model 4 participants created educational programs and materials for their patients. These included the following: spine surgery classes, joint surgery classes, and informational pamphlets. One Awardee described that, prior to their redesign, their patients received information from multiple sources in the hospital, which were not always consistent. After joining BPCI, they created a patient guidebook that maps out the process patients should expect during their care, describes the patient's own role in getting better, and sets expectations with regard to length of stay. Several participants identified patient engagement as a key factor to their success in the initiative.

> *"We've had fairly significant success in reducing implant costs and we've also had good success in the project of reducing our readmissions in some of our hospital acquired conditions, so those have all resulted in cost savings."*

All of the Model 4 case study and quarterly call participants described standardizing their care protocols as an element of care redesign. For example, one Awardee discovered that its surgeons were each using different pain management protocols; as a result of BPCI, they streamlined these protocols and stopped using a resource-intensive procedure that did not produce better clinical outcomes. Other participants standardized the devices used by surgeons so that they could negotiate better rates from vendors. Some Model 4 sites also developed detailed discharge checklists that are sent to SNFs that receive BPCI beneficiaries and they have developed detailed guides for how and when patients should be contacted by hospital staff (e.g., days 1-7 post surgery, then again at 10, 15, 20, 25, etc.). Though some Model 4 participants implemented care redesign initiatives exclusively for their BPCI beneficiaries, one site noted that any changes in care would be extended to all patients at their facility.

During quarterly call and case study interviews, Model 4 participants described the use of case managers/care navigators to reduce readmissions. Though each participant assigned slightly different functions to these positions, these staff members generally establish a relationship with the patient prior to surgery, discuss options for care after discharge, conduct follow-up calls

throughout the 30-day episode, and teach pre-surgery classes. Some Awardees employed multiple staff in these new positions, whereas others had only one individual doing this work.

According to some Model 4 participants, tracking readmissions to non-affiliated hospitals was challenging. These participants reported that the data they rely on for information on readmissions

> *"[Our navigator] goes beyond the hospital walls. On the pre-op side, she meets patients about two weeks before they're admitted. She teaches a class and establishes a relationship with them, and they are in contact with her sometimes before the admission, so she definitely is on the front end. On the back end, she's been doing phone calls, follow-up phone calls, and checking in with them if they go to skilled nursing."*

was incomplete; it often indicated that a beneficiary was readmitted but did not explain the cause of the readmission. In addition, participants noted that there was a delay in receiving this readmission data from CMS. These challenges related to tracking beneficiary data were reported to inhibit efforts to improve readmission rates. These concerns were also raised during interviews with three Model 4 participants that withdrew from the program as of June 2015. All of these participants indicated that they faced significant challenges in utilizing the data files that they received from CMS, which in turn led to delayed payments to physicians. This was the key reason that the Model 4 participants whom we spoke with chose to discontinue participation in BPCI.

During site visits, some EIs also indicated challenges associated with staffing that have resulted from their participation in BPCI. For example, participants reported a need for additional full time employees, increased staff responsibilities for existing staff members, and additional training for staff members to understand the goals of BPCI. One participant reported that it was "difficult for some cardiologists to conceptualize the incentive." This was noted as a challenge because promoting changes under BPCI required the participants to incentivize their physicians. Another Awardee described challenges associated with the turnover of hospital leadership, as the incoming leadership prioritized a new agenda that did not focus on BPCI to the extent the hospital had in the past.

b. Cost savings strategies

The primary cost saving strategy described by all Model 4 participants was reducing readmissions. On site interviews and quarterly calls, we heard of different approaches to achieving this goal. One Awardee indicated that surgeons have taken increased ownership of post-discharge care to prevent unnecessary readmissions from the ER. Other sites relied on the efforts of case managers/care navigators to schedule follow-up appointments and identify high-risk beneficiaries. Several participants stated that they saw a reduction in readmission rates since they began participating in BPCI. Another Awardee had no readmissions for their BPCI beneficiaries, but two non-BPCI patients were readmitted, which led the Awardee to adopt the care redesign practices for all its patients.

Care standardization was another strategy participants implemented in an effort to decrease costs. Several Model 4 participants reported, however, that it was challenging to introduce new protocols and to standardize care when physicians are accustomed to their own methods of providing care. Describing this situation, one Awardee stated simply, "Physicians don't want to change the way they have always practiced." Nevertheless, Model 4 participants still identified adherence to new care protocols as a key determinant of cost savings under BPCI.

B. Impact of BPCI

This section presents the first year impact estimates of Model 4 BPCI (hospital) episodes on payments, utilization, and quality of care based on episodes that were initiated during the first four quarters of the BPCI initiative (Q4 2013 through Q3 2014).[76]

We present results for the key outcomes across all Model 4 episodes here and then separately for the two Model 4 clinical episode groups with at least 10 BPCI hospital EIs, i.e. orthopedic surgery (N=14) and cardiovascular surgery (N=10). The exhibits present the estimated differential change in risk-adjusted outcomes for beneficiaries receiving care from BPCI providers between the baseline and the intervention period relative to the same change for the beneficiaries receiving care from providers in a comparison group (DiD). See Section II.D.2 for additional details on the statistical approach.

1. Change in Medicare standardized allowed payment amounts

We calculated the Medicare standardized allowed payment amount by type of service as well as across all services for two measurement periods. The first measurement period is relative to the length of the episode (30 days). We present the total payments included in the bundle and the total payments not included in the bundle. We also show results for total payment outcomes for the 60 days after the end of the episode (post-bundle period) and for the 30 days prior to the start of the episode (pre-bundle period.) The second measurement period is the anchor hospitalization and the 90 days after the hospital discharge (90-day PDP). This section summarizes the results for key payment outcomes.

Exhibit 142 presents the direction and point estimate of the DiD estimate for key payment outcomes and the two clinical episode groups with sufficient EI participation for Model 4. The top panel shows payment outcomes that were measured based on episodes initiated during the first year of the initiative. The bottom panel shows payment outcomes that were measured based on episodes initiated during the first three quarters of the initiative to allow for sufficient claims run out. These include post-bundle and pre-bundle payment outcomes. Orange highlighted cells represent a statistically significant decrease relative to the comparison group. Green highlighted cells represent a statistically significant increase relative to the comparison group. The baseline and intervention estimates, for both BPCI and the comparison group, can be located for key outcomes in Sections V.D through V.E.

Overall, the change in standardized allowed payment amounts from baseline to intervention period for Model 4 BPCI hospital EIs was not statistically significant relative to the change for beneficiaries treated by the matched comparison group of providers. For BPCI cardiovascular surgery episodes, although the relative decline in standardized allowed amount for the bundle was not statistically significant, there was a statistically significant relative increase in standardized allowed payment amounts for services not included in the bundle definition during the 30 days post bundle.

---

[76] Please note that pre-bundle and post bundle period (PBP) payment outcomes and patient-assessment based outcomes are reported with a one-quarter delay and therefore are based on episodes that were initiated during the first three quarters of the BPCI initiative (Q4 2013 through Q2 2014).

Exhibit 142: Diff-in-Diff Estimate for Standardized Allowed Payment Amounts, by Clinical Episode Group, Model 4, Baseline to Intervention

Measure	Orthopedic Surgery	Cardiovascular Surgery
Number of episodes initiated Q4 2013 – Q3 2014	1,848	1,173
BPCI Cumulative to Date (Q4 2013 – Q3 2014)		
Standardized allowed amount in bundle definition	$92	-$789
Standardized allowed amount (Part A & B), IP through 90-day post-discharge period	$464	-$1,100
BPCI First Three Quarters (Q4 2013 – Q2 2014)		
Other standardized allowed amount Part A & B, days 1-30 Post-bundle Payment	-$144	$396
Other standardized allowed amount Part A & B, days 31-60 Post-bundle Payment	-$11	$0
Standardized allowed amount Part B, 30-day pre-bundle payment	$64	$71

Note: Statistical significance at the 0.05 level is indicated by orange and green shaded cells. Orange indicates the DiD estimate was negative and statistically significant; green indicates the DiD estimate was positive and statistically significant. The results presented in this table represent episodes initiated in the first four quarters of the BPCI Initiative (Q4 2013 through Q3 2014). Please note that pre-bundle and post bundle period (PBP) payment outcomes are reported with a one-quarter delay.

Source: Lewin analysis of Medicare claims and enrollment data for episodes that began Q4 2011 through Q3 2014 for BPCI providers and a comparison group.

We calculated a number of utilization outcomes on institutional PAC and HH use to understand the changes that may have contributed to the payment outcomes summarized above. Exhibit 143 shows key utilization measures. The baseline and intervention estimates, for both BPCI and the comparison group, can be located for key outcomes in Sections IV.D through IV.I.

Although not statistically significant at the 0.05 level, the relative increase in cardiovascular surgery patients discharged to institutional PAC likely contributed to the higher post-bundle payments. Furthermore, the 1.9 day increase in home health visits for home health users was statistically significant and the other PAC use measures suggest increased reliance on PAC. These factors may have contributed to lower payments in the bundle, particularly if PAC use was associated with lower readmissions. Though BPCI beneficiaries in orthopedic surgery episodes who used PAC exhibited a significant decline relative to comparison beneficiaries in the total number of days in any institutional setting through the 90-day PDP, this was not associated with a statistically significant decrease in total payment for Part A and B services during the anchor hospitalization and the 90-day PDP. This may be because there may have been more orthopedic surgery patients discharged to PAC, offsetting the lower costs associated with reduced institutional days. Similarly, among beneficiaries treated for cardiovascular surgery episodes who were discharged to HH, beneficiaries treated by BPCI providers received more HH visits in the 90-day post-discharge period relative to the comparison group (1.9 visits), but there was no statistically significant change in total payment for Part A and B services during the anchor hospitalization and the 90-day PDP from baseline to intervention.

Exhibit 143: Diff-in-Diff Estimate for Utilization Outcomes, by Clinical Episode Group, Model 4, Baseline to Intervention

Measure	Orthopedic Surgery	Cardiovascular Surgery
Number of episodes initiated Q4 2013 – Q3 2014	1,848	1,173
% discharged to an institution out of those who received any post-acute care	0.8 pp	5.7 pp
% discharged to post-acute care	4.5 pp	-0.1 pp
Acute Inpatient Care LOS	-0.2	-0.1
Number of HH Visits, 90-day PDP[1]	0.7	1.9
Number of Institutional Days, 90-day PDP[1]	-3.0	1.6
Number of IRF Days, 90-day PDP[1]	0.0	0.0
Number of Readmission Days, 90-day PDP[1]	-0.4	-0.4
Number of SNF Days, 90-day PDP[1]	-2.0	2.2

[1] The number of visits/days is conditional on having at least 1 day in the respective setting.

Note: Statistical significance at the 0.05 level is indicated by orange and green shaded cells. Orange indicates the DiD estimate was negative and statistically significant; green indicates the DiD estimate was positive and statistically significant. The results presented in this table represent episodes initiated in the first four quarters of the BPCI Initiative (Q4 2013 through Q3 2014). A blank cell indicates that the outcome cannot be presented, either due to insufficient sample size or the type of episodes initiated during the time period. pp=percentage points.

Source: Lewin analysis of Medicare claims and enrollment data for episodes that began Q4 2011 through Q3 2014 for BPCI providers and a comparison group.

2. *Changes in beneficiary quality of care*

Exhibit 144 presents the impact of Model 4 BPCI on key claims-based and beneficiary assessment-based quality of care measures by clinical episode group. The outcomes are stratified in the Exhibit based on the quarters included in the DiD estimate; claim-based outcomes were measured based on episodes initiated during the first year of the initiative, while assessment-based outcomes were measured based on episodes initiated during the first three quarters of the initiative in order to allow for sufficient assessment data run out. Orange highlighted cells represent a statistically significant decrease relative to the comparison group. Green highlighted cells represent a statistically significant increase relative to the comparison group.

The claims-based measures of quality of care include mortality within 30 days of discharge as well as emergency department use and readmission rates during the 90 days after discharge from the anchor hospitalization. The beneficiary assessment outcomes include improvement or change in functional status among beneficiaries who received post-acute care. See **Appendix L** for detailed outcome definitions.   The baseline and intervention estimates, for both BPCI and the comparison group, can be located for key outcomes in Sections IV.D through IV.I

a. *Claim-based measures*

Although the point estimates for claims-based measures indicate declines, there were no statistically significant changes between baseline and intervention periods in 30-day mortality, 90-day emergency department visits, or 90-day readmission rates for either the orthopedic surgery or cardiovascular surgery clinical groups among beneficiaries treated by BPCI providers relative to beneficiaries treated by matched comparison providers.

b. Assessment based measures

There were no statistically significant differences in the changes in assessment-based quality measures from baseline to intervention period between Model 4 BPCI beneficiaries and comparison beneficiaries who received their first PAC treatment at a SNF or IRF. For BPCI beneficiaries with cardiovascular surgery episodes who received their first PAC treatment at a HHA, however, the changes in share of beneficiaries that exhibited improvement in bathing, ambulation, and bed transferring declined from baseline to intervention period relative to the comparison beneficiaries. For these three functional measures, the share of beneficiaries that exhibited improvement declined among the BPCI Model 4 beneficiaries and increased among the comparison group beneficiaries from baseline to intervention period, resulting in a negative effect of BPCI intervention (9.2, 10.9 and 14.8 percentage points for bathing, ambulation and bed transferring, respectively). This occurred despite the relative increase in home health visits noted above.

Exhibit 144: Diff-in-Diff Estimate for Claim-based and Assessment-based Quality Outcomes, by Clinical Episode Group, Model 4, Baseline to Intervention

Measure	Orthopedic Surgery	Cardiovascular Surgery
Number of episodes initiated Q4 2013 – Q3 2014	1,848	1,173
BPCI Cumulative to Date (Q4 2013 – Q3 2014)		
All-cause mortality rate, 30-day PDP	-0.2 pp	-0.6 pp
Emergency Department Use, 30-day PDP	-1.3 pp	-2.3 pp
Emergency Department Use, 90-day PDP	-0.7 pp	-2.3 pp
Unplanned readmission rate, 30-day PDP	-0.9 pp	-3.1 pp
Unplanned readmission rate, 90-day PDP	-2.2 pp	-3.6 pp
BPCI First Three Quarters (Q4 2013 – Q2 2014)		
ADL HHA, improved ambulation	3.5 pp	-10.9 pp
ADL HHA, improved bathing	1.7 pp	-9.2 pp
ADL HHA, improved bed transferring	3.9 pp	-14.8 pp
ADL HHA, improved lower body dressing	-0.6 pp	2.6 pp
ADL HHA, improved upper body dressing	3.1 pp	-3.0 pp
ADL IRF, average change in mobility score	1.2 pp	0.1 pp
ADL IRF, average change in self-care score	1.1 pp	0.7 pp
ADL SNF, improved mobility function	-0.5 pp	13.9 pp
ADL SNF, improved overall function	1.3 pp	-3.3 pp
ADL SNF, improved self-care function	1.2 pp	0.9 pp

Note: Statistical significance at the 0.05 level is indicated by orange and green shaded cells. Orange indicates the DiD estimate was negative and statistically significant; green indicates the DiD estimate was positive and statistically significant. The results presented in this table represent episodes initiated in the first four quarters of the BPCI Initiative (Q4 2013 through Q3 2014). Please note that assessment-based quality measures are reported with a one-quarter delay. A blank cell indicates that the outcome cannot be presented, either due to insufficient sample size or the type of episodes initiated during the time period. pp=percentage points.

Source: Lewin analysis of Medicare claims and enrollment data for episodes that began Q4 2011 through Q3 2014 for BPCI providers and a comparison group.

C.   Market Dynamics

For Model 4 hospital EIs, we calculated the market shares of MJRLE, CHF, and sepsis episodes to test our hypothesis that hospital EIs would strive to increase their market share of BPCI-eligible episodes.[77]   The table in the section below shows trends for the MJRLE market share of all Model 4 EIs across national markets.   The same tables for CHF and sepsis episodes can be found in **Appendix P**. We also present findings for Model 4 EIs in two individual markets we studied: Orlando, FL (Orlando market) and Kalamazoo-Portage, MI (Kalamazoo market). Given the small counts of EIs and the preliminary nature of these analyses, some market effects may have gone undetected.

1.   Market share of Model 4 EIs

Exhibit 145 shows the trends for the market share of MJRLE episodes across all Model 4 EIs and the average number of episodes per EI. Median market shares were substantially lower than the mean market shares across Model 4 EIs, but both trends indicate that Model 4 EIs increased their market share of BPCI-eligible MJRLE episodes over time. There is some indication of an upward trend starting before the beginning of the BPCI program. For BPCI-eligible MJRLE episodes, the average market share of Model 4 EIs increased from 17-18% in the pre-BPCI period to 19.3% in the BPCI period. The median increased from 5% before to 7.1% after BPCI.

Out of 20 BPCI EIs participating in Model 4, one hospital EI was located in the Orlando market and one was located in the Kalamazoo market (Exhibit 145). Consistent with our hypothesis that EIs will try to capture greater market shares of BPCI-eligible episodes, the single Model 4 EI in each of these markets had their highest market shares of MJRLE episodes in the two BPCI intervention periods.[78] The Orlando EI had market shares of 46.3% and 45.7% in the two BPCI intervention periods, compared with the next highest market share of 43.3% in Q2 2012/ Q3 2012 pre-BPCI period. The Kalamazoo EI had market shares of 53.0% and 55.3% in the two BPCI intervention periods, compared with the next highest market share of 49.9% percent in Q2 2013/ Q3 2013 in the pre-BPCI period.

The trend in market shares for all Model 4 markets combined for sepsis episodes was similar to the trend for MJRLE (**Appendix P**). The average market share among all Model 4 EIs increased steadily from around 19% to over 20% in the last pre-BPCI period and remained over 20% in the one year intervention period. The median increased from 7.2% in the first pre-BPCI period to 11.1% in the one year intervention period. The average market share of BPCI-eligible CHF episodes showed no change over time, but the median market share of CHF episodes was 12.3% in Q2 2014/ Q3 2014 and the highest median market share in the pre-BPCI period was 11.3%, indicating an upward trend in the median market share (**Appendix P**). There was little change over time in the Orlando or Kalamazoo EI's market shares for CHF or sepsis episodes.

---

[77] The market share of hospital EIs is the number of BPCI-eligible episodes admitted to the EI, divided by the total number of the same type of episodes admitted to all hospitals in the market (defined as the CBSA in which the EI is located).

[78] We did not estimate statistical significance of the difference in measure rates between the baseline and intervention period due to small sample size. Even at the national level (all EIs across all markets), the power analysis suggested the sample size was too small to reliably detect a statistically significant difference if one was present. We will re-evaluate this next year.

Exhibit 145: Average Market Share of Episode Initiators for Major Joint Replacement of the Lower Extremity Episodes, Model 4, Q2 2011-Q3 2014

	Baseline				Intervention	
	Q4 2011 / Q1 2012	Q2 2012 / Q3 2012	Q4 2012 / Q1 2013	Q2 2013 / Q3 2013	Q4 2013 / Q1 2014*	Q2 2014 / Q3 2014
ALL MARKETS (20 EIs)						
EI Market Share						
Mean	17.7%	17.2%	17.3%	18.8%	19.2%	19.3%
SD	24.0%	22.4%	23.0%	24.6%	25.2%	24.8%
Median	5.0%	6.8%	6.3%	7.0%	7.1%	7.1%
25th Pctl	2.8%	3.0%	2.7%	2.9%	3.0%	3.2%
75th Pctl	27.5%	26.4%	26.5%	27.6%	29.6%	30.6%
Number of MJRLE episodes per EI						
Mean	131.2	125.4	136.6	140.9	159.0	151.1
ORLANDO MARKET (1 EI)						
EI Market Share						
Mean	42.6%	43.3%	41.4%	42.6%	46.3%	45.7%
Number of MJRLE episodes per EI						
Mean	696	634	690	658	852	767
Number of hospitals with MJRLE admissions in the market						
Total	9	9	9	9	9	9
KALAMAZOO MARKET (1 EI)						
EI Market Share						
Mean	46.5%	48.9%	46.2%	49.9%	53.0%	55.3%
Number of MJRLE episodes per EI						
Mean	153	160	160	181	205	197
Number of hospitals with MJRLE admissions in the market						
Total	3	3	3	3	3	3

* One Model 4 hospital EI joined BPCI in Q4 2013. All others joined in Q1 2014. *Orlando* and *Kalamazoo*: Both Model 4 hospital EIs joined BPCI in Q1 2014.

SD = standard deviation

Source: Lewin analysis of Medicare claims and enrollment data for episodes that began Q4 2011 through Q3 2014 for BPCI participants and a comparison group.

D.   Impact of BPCI on Orthopedic Surgery

1.   *Participants*

During the first four quarters of the initiative, 70% of EIs in Model 4 (14 hospitals) participated in at least one orthopedic surgery episode. There were 1,848 orthopedic surgery episodes initiated in Model 4-participating hospitals, accounting for 50% of all Model 4 episodes. Nearly all (96%) of these episodes were for major joint replacement of the lower extremity MS-DRGs. All Model 4 hospitals participating in BPCI with these clinical episodes were urban facilities and 64% were non-profits. They were larger than the typical hospital (with an average of 331 beds, compared with 188 beds in non-participating hospitals) and likelier to be teaching hospitals.

2.   Patient population characteristics

As shown in Exhibit 146, beneficiaries who were treated in a hospital participating in Model 4 for an orthopedic surgery episode were similar to all Medicare beneficiaries treated in the same MS-DRG with respect to gender and dual eligibility status. The major differences between the BPCI and non-BPCI groups were related to age and Medicare eligibility due to disability. The beneficiaries of BPCI providers were older than all beneficiaries with the same MS-DRG, with a larger proportion aged 65-79 (64.8% vs. 58.9%) and a smaller proportion aged 20-64 (6.4% vs. 11.5%). Among BPCI beneficiaries, 8.8% were eligible for Medicare due to disability, compared with 13.3% of all Medicare beneficiaries.

Exhibit 146: Characteristics of BPCI Beneficiaries and All Medicare Beneficiaries with a Hospitalization for Orthopedic Surgery MS-DRG, Model 4, Q4 2013 - Q3 2014

Characteristics	Model 4 Orthopedic Surgery Episodes BPCI Beneficiaries, Q4 2013 – Q3 2014 (N = 1,749)		All Medicare Beneficiaries With Same MS-DRG Admission, Q3 2014 (N = 196,694)	
	N	%	N	%
Age				
20-64	112	6.4	22,694	11.5
65-79	1,134	64.8	115,834	58.9
80+	503	28.8	58,166	29.6
Gender				
Female	1,159	66.3	127,405	64.8
Male	590	33.7	69,289	35.2
Medicaid and Disability				
Eligible Medicaid	263	15.0	29,707	15.1
Disability, no ESRD	154	8.8	26,129	13.3

Source: Lewin analysis of Medicare claims and enrollment data for episodes that began Q4 2011 through Q3 2014 for BPCI providers and for all admissions for the same MS-DRG to non-participating hospitals in Q3 2014.

3.   Change in Medicare standardized allowed payments

There were no statistically significant differences between Model 4 BPCI episodes and episodes initiated by matched providers from the baseline to the intervention periods in Medicare standardized allowed payments. The average total standardized allowed payment for Part A and B services during the anchor hospitalization and the 90-day PDP for patients treated by BPCI providers increased from baseline to intervention periods, but this increase was not statistically significant relative to the change for beneficiaries treated by matched comparison providers (see Exhibit 147). Furthermore, there were no statistically significant differences in the change for the allowed amount for any Part A service (i.e. SNF, IRF). BPCI providers, however, did have a greater decline in number of institutional days relative to the comparison group providers (-3 days) that was statistically significant at 0.05. For other forms of PAC use, there were no statistically significant differences in the change – in either the first setting or the intensity of services (i.e. length of stay or number of home health visits) - relative to the episodes from matched comparison providers. There were no indications that these results were achieved by shifting services outside of the bundle definition or bundle period.

Exhibit 147: Impact of BPCI on Total Standardized Allowed Payment Inpatient Anchor Hospitalization and 90-day PDP for Orthopedic Surgery Episodes, Model 4, Q4 2011-Q3 2014

	BPCI (N=1,708)		Comparison (N=1,710)		Diff-in-Diff estimate		
Measure	Baseline	Intervention	Baseline	Intervention	Diff-in-Diff	LCI	UCI
Total Standardized Allowed Payment, Inpatient Anchor Hospitalization and 90-day PDP	$28,079	$28,157	$29,574	$29,188	$464	-$1,181	$2,109

* Denotes statistical significance at the 5% level.

LCI = lower confidence interval at the 5% level; UCI = upper confidence interval at the 5% level.

Note: Baseline is defined as episodes that began Q4 2011 through Q3 2012. Intervention is defined as episodes that began Q4 2013 through Q3 2014. Sample sizes reflect number of episodes initiated during the intervention period that met inclusion criteria for the outcome.

Source: Lewin analysis of Medicare claims and enrollment data for episodes that began Q4 2011 through Q3 2014 for BPCI and comparison providers.

4. Change in beneficiary quality of care

The change in quality of care for BPCI beneficiaries with orthopedic surgery episodes was not statistically different from that of beneficiaries of non-BPCI providers, as measured through various claim-based measures and PAC assessments. Hospital readmission rates, emergency department visits, and mortality following the anchor hospitalization for orthopedic surgery episodes all declined for BPCI episodes relative to the comparison episodes, but these changes were not statistically different. There were no statistically significant changes in assessment-based quality measures from baseline to intervention period between BPCI and comparison group beneficiaries with orthopedic surgery episodes who received their first PAC treatment at a SNF, an HHA or an IRF.

Exhibit 148: Impact of BPCI on Claims-based Quality Measures for Orthopedic Surgery Episodes, Model 4, Q 2011-Q3 2014

	BPCI (N=1,729)		Comparison (N=1,739)		Diff-in-Diff estimate		
Measure	Baseline	Intervention	Baseline	Intervention	Diff-in-Diff	LCI	UCI
Unplanned Readmission Rate, within 30 days	6.4%	4.7%	7.0%	6.2%	-0.9	-3.0	1.1
Unplanned Readmission Rate, within 90 days	10.2%	8.5%	11.1%	11.7%	-2.2	-5.1	0.7
Emergency Department Visit Rate, within 30 days	7.4%	6.5%	6.8%	7.3%	-1.3	-3.8	1.2
Emergency Department Visit Rate, within 90 days	12.5%	12.7%	12.3%	13.2%	-0.7	-3.8	2.5
Mortality, within 30 days	1.2%	0.8%	1.1%	0.9%	-0.2	-1.2	0.8

* Denotes statistical significance at the 5% level.

LCI = lower confidence interval at the 5% level; UCI = upper confidence interval at the 5% level.

Note: Baseline is defined as episodes that began Q4 2011 through Q3 2012. Intervention is defined as episodes that began Q4 2013 through Q3 2014. Sample sizes reflect number of episodes initiated during the intervention period that met inclusion criteria for the outcome.

Source: Lewin analysis of Medicare claims and enrollment data for episodes that began Q4 2011 through Q3 2014 for BPCI and comparison providers.

E. Impact of BPCI on Cardiovascular Surgery

1. Participants

During the first four quarters of the initiative, 10 hospitals, or 50% of hospital EIs in Model 4, participated in at least one cardiovascular surgery episode.[79] There were 1,173 cardiovascular surgery episodes initiated in Model 4-participating hospitals, accounting for 31% of all Model 4 episodes. Approximately 34% of these episodes were for percutaneous coronary intervention MS-DRGs and 30% were for coronary artery bypass MS-DRGs. Each of the hospitals participating in BPCI with these clinical episodes was an urban facility and 60% were non-profits. They were larger than the typical hospital (with an average of 511 beds, compared with 188 beds in non-participating hospitals) and more likely to be teaching hospitals.

2. Patient population characteristics

Beneficiaries who were treated in a hospital participating in Model 4 for a cardiovascular surgery episode were different from all Medicare beneficiaries treated in the same MS-DRG. The major differences between BPCI and non-BPCI groups were related to gender and Medicaid eligibility. A larger share of beneficiaries of BPCI providers were male than all beneficiaries with the same MS-DRG (37.0% vs. 42.1%; Exhibit 149). Among BPCI patients, 10.6% were eligible for Medicaid, compared to 17.9% of all Medicare beneficiaries. Both groups were similar with respect to age and Medicare eligibility due to disability.

Exhibit 149: Characteristics of BPCI Beneficiaries and All Medicare Beneficiaries with a Hospitalization for a Cardiovascular Surgery MS-DRG, Model 4, Q4 2013 - Q3 2014

Characteristics	Model 4 Cardiovascular Surgery Episodes BPCI Beneficiaries, Q4 2013 – Q3 2014 (N = 1,118)		All Medicare Beneficiaries With Same MS-DRG Admission, Q3 2014 (N = 62,576)	
	N	%	N	%
Age				
20-64	91	8.1	8,023	12.8
65-79	646	57.8	34,213	54.7
80+	381	34.1	20,340	32.5
Gender				
Female	414	37.0	26,340	42.1
Male	704	63.0	36,236	57.9
Medicaid and Disability				
% Eligible Medicaid	119	10.6	11,177	17.9
% Disability, no ESRD	118	10.6	6,931	11.1

Source: Lewin analysis of Medicare claims and enrollment data for episodes that began Q4 2011 through Q3 2014 for BPCI providers and for all admissions for the same MS-DRG to non-participating hospitals in Q3 2014.

---

[79] The clinical episodes that compose the cardiovascular surgery group are: AICD generator or lead; cardiac defibrillator; cardiac valve; coronary artery bypass graft; major cardiovascular procedure; other vascular surgery; pacemaker; pacemaker device replacement or revision; and percutaneous coronary intervention.

3. Change in Medicare standardized allowed payment amounts

The $1,100 relative decline in total standardized allowed amounts for the anchor hospitalization plus 90-day PDP for BPCI episodes was not statistically significant relative to the change for the comparison group (see Exhibit 150). There was a statistically significant increase in the number of home health visits relative to the comparison group of 1.9 days. However, payments for services not included in the bundle definition increased $396 more for BPCI than comparison episodes during the 30 day post bundle period. This statistically significant relative increase may be associated with the decline in payments for the anchor hospitalization and 90-day PDP.

Exhibit 150: Impact of BPCI on Total Standardized Allowed Payment Inpatient Anchor Hospitalization and 90-day PDP for Cardiovascular Surgery Episodes, Model 4, Q4 2011-Q3 2014

Measure	BPCI (N=1,006)		Comparison (N=1,074)		Diff-in-Diff estimate		
	Baseline	Intervention	Baseline	Intervention	Diff-in-Diff	LCI	UCI
Total Standardized Allowed Payment, Inpatient Anchor Hospitalization and 90-day PDP	$40,088	$39,711	$39,288	$40,012	-$1,100	-$3,940	$1,739

* Denotes statistical significance at the 5% level.

LCI = lower confidence interval at the 5% level; UCI = upper confidence interval at the 5% level.

Note: Baseline is defined as episodes that began Q4 2011 through Q3 2012. Intervention is defined as episodes that began Q4 2013 through Q3 2014. Sample sizes reflect number of episodes initiated during the intervention period that met inclusion criteria for the outcome.

Source: Lewin analysis of Medicare claims and enrollment data for episodes that began Q4 2011 through Q3 2014 for BPCI and comparison providers.

4. Change in beneficiary quality of care

The change in quality of care for BPCI beneficiaries with cardiovascular surgery episodes was not statistically significantly different from that of beneficiaries of non-BPCI providers, as measured through various claims-based measures, although all six measures indicated relative improvements in quality (Exhibit 151).

Exhibit 151: Impact of BPCI on Claims-based Quality Measures for Cardiovascular Surgery Episodes, Model 4, Q4 2011-Q3 2014

Measure	BPCI (N=1,087)		Comparison (N=1,087)		Diff-in-Diff estimate		
	Baseline	Intervention	Baseline	Intervention	Diff-in-Diff	LCI	UCI
Unplanned Readmission Rate, within 30 days	17.2%	12.4%	13.1%	11.4%	-3.1	-7.0	0.8
Unplanned Readmission Rate, within 90 days	27.6%	23.3%	22.6%	21.9%	-3.6	-8.0	0.9
Emergency Department Visit Rate, within 30 days	10.2%	10.7%	12.3%	15.0%	-2.3	-5.4	0.8
Emergency Department Visit Rate, within 90 days	19.2%	19.1%	22.3%	24.5%	-2.3	-5.8	1.2
Mortality, within 30 days	1.9%	1.8%	1.5%	2.0%	-0.6	-1.6	0.5

* Denotes statistical significance at the 5% level.

LCI = lower confidence interval at the 5% level; UCI = upper confidence interval at the 5% level.

Note: Baseline is defined as episodes that began Q4 2011 through Q3 2012. Intervention is defined as episodes that began Q4 2013 through Q3 2014.

Source: Lewin analysis of Medicare claims and enrollment data for episodes that began Q4 2011 through Q3 2014 for BPCI and comparison providers.

There were statistically significant changes in the share of beneficiaries that exhibited improvement in bathing, ambulation and bed transferring between the BPCI beneficiaries and comparison group patients in Model 4 with cardiovascular surgery episodes who received their first PAC at a HHA (Exhibit 152). For all three measures, the share of beneficiaries that exhibited improvement declined from baseline to intervention period among BPCI beneficiaries and increased among the comparison group beneficiaries, resulting in a statistically significant and negative effect of the BPCI intervention on these measures. For example, the share of beneficiaries that exhibited improvement in bathing dropped 2.4 percentage points among the BPCI beneficiaries and increased by 6.8 percentage points among comparison beneficiaries, which led to a statistically significant and negative effect of BPCI intervention (9.2 percentage points).

Exhibit 152: Impact of BPCI on Assessment-based Quality Measures for Cardiovascular Surgery Episodes, Model 4, Q4 2011 – Q2 2014

Measure	BPCI (N=132)		Comparison (N=126)		Diff-in-Diff estimate		
	Baseline	Intervention	Baseline	Intervention	Diff-in-Diff	LCI	UCI
Percent of HHA beneficiaries with improved bathing	84.3%	81.9%	83.8%	90.6%	-9.2 *	-15.4	-2.9
Percent of HHA beneficiaries with improved ambulation	82.0%	77.0%	80.5%	86.4%	-10.9 *	-19.9	-2.0
Percent of HHA beneficiaries with improved bed transferring	84.5%	73.3%	80.6%	84.2%	-14.8 *	-28.4	-1.2

* Denotes statistical significance at the 5% level.

LCI = lower confidence interval at the 5% level; UCI = upper confidence interval at the 5% level.

Note: Baseline is defined as episodes that began Q4 2011 through Q3 2012. Intervention is defined as episodes that began Q4 2013 through Q2 2014. Sample sizes reflect number of episodes initiated during the intervention period that met inclusion criteria for the outcome.

Source: Lewin analysis of patient assessment data for episodes that began Q4 2013 through Q2 2014 for BPCI and comparison providers.

VI. Discussion

This second Annual Report provides a summative and formative evaluation of the BPCI initiative based on quantitative analyses of Phase 2 participants that joined the initiative during the first year and qualitative analyses of participants that joined during the first seven quarters. We present information about the providers that chose to participate in the initiative and how they differ from other providers. We summarize the participants' various approaches to implementing BPCI, including their selection of Model and clinical episodes, relationships with partners, use of waivers, and care redesign and cost saving strategies. We analyze changes in key outcomes from before BPCI to after implementation for BPCI episodes and comparable episodes from comparison providers. This report updates and expands the information contained in the first Annual Report, which was restricted to the earliest BPCI participants and their initial experiences during the first quarter of the initiative. Not only does this Report include a broader range of participants and longer exposure to the BPCI incentives, but it contains new data from the beneficiary survey on satisfaction, health care experience, and changes in functional status for various subgroups of beneficiaries in BPCI episodes. We also introduce additional analyses to address questions about market dynamics in the communities where BPCI participants are located and factors that distinguish BPCI participants that achieved favorable payment and quality outcomes from those that did not.

Most results are based on the experience of 94 Awardees across three Models, with 227 episode initiators that were responsible for 58,410 episodes of care during the first year of the initiative. Because BPCI participants could choose among Models, episodes, and waivers, our analyses stratify outcomes in multiple ways. We remain limited in our ability to estimate the impact of the initiative under most Model and episode combinations, however, because of insufficient sample size and the limited time the initiative has been underway. This report focuses on the situations in which there were statistically significant results for key outcome measures and consistent evidence from qualitative sources.

A.   Range of Choices and Impact on Results

The BPCI initiative was designed so that participants had multiple implementation options, including the fundamental choice of whether to participate or not. Descriptive statistics reveal that providers that chose to participate in Phase 2 during the first year differ from those that delayed participation or chose not to participate. The hospitals, SNFs, and HHAs that participated in BPCI as EIs were more likely to be larger, urban facilities, generally located in areas with higher income populations than other providers of the same type. Episodes from BPCI-participating facilities also had higher standardized Medicare allowed Part A payments prior to joining BPCI. Taken together with information from participant interviews about preparations for entering the initiative and participants' on-going change efforts, these characteristics indicate that participants may have more financial or leadership resources to devote to responding to changing payment incentives.

During preparations for BPCI prior to Phase 2, many participants told us that consultants advised them on choices about how to participate in BPCI, including gainsharing methods, and provided on-going assistance with monitoring performance and analyzing CMS-provided data. The availability of financial resources to obtain this outside help may be an important distinguishing factor between those that participated and those that did not. In the future we will explore whether Awardees or Conveners provide the resources EIs require to participate successfully.

Experience to date indicates that the Convener approaches are quite prevalent in the participation of PAC providers under Model 3.

Beyond the decision on whether or not to participate, participants must select their BPCI Model and episodes. Hospital EIs indicated that these decisions were based on opportunities to reduce spending and perceived risk. Often hospitals chose Model 2 because they believed they could reduce PAC use; those that entered under Model 4 indicated a lack of ability to influence PAC. As to risk, many hospital EIs we spoke with indicated that they considered the financial risk of their Model choice. Given the much larger participation in Model 2, it must have been considered financially favorable or less risky or both. For hospitals, episode choice was often determined by whether they had sufficient volume of cases within the episode type and the potential for savings. Many of them mentioned a physician champion in a particular specialty that affected their choice. PAC EIs under Model 3 generally chose more episode groups, possibly because of their overall smaller size. Their choice of episodes may have been influenced by the Awardee Convener most of them participated with.

B.   Impact of BPCI on Costs and Quality

For most clinical episodes, there were no statistically significant differences in the change in Medicare standardized allowed payments between BPCI participants and comparison providers, although many of the participants we interviewed indicated that they had implemented efforts intended to reduce total episode costs. The lack of widespread impact of BPCI on Medicare payments across clinical episodes thus far may be due to several factors. The quantitative results are based on less than one full year of BPCI experience for the majority of participants. Although some providers have been preparing for the initiative well before entering into Phase 2, this still may not be enough time to see the impact of these changes on episode payments. There is limited sample size for most clinical episodes, which limits our ability to detect changes in payments. Further, to increase our ability to detect changes, we conducted many of our analyses on aggregations of clinical episodes, which may mask payment changes for particular types of episodes.

For Model 2 orthopedic and cardiovascular surgery episodes participants' efforts to reduce episode spending are achieving expected results. For these episodes, which account for a large share of Model 2 episodes, we saw a statistically significant shift from more expensive institutional PAC to less expensive home health care among beneficiaries discharged to any PAC setting. This shift was the major contributor to the larger relative decline in total payments during the anchor stay and the 90-day PDP for orthopedic surgery episodes. There was not, however, a statistically significant decline in payments during the anchor stay and the 90-day PDP for cardiovascular surgery episodes. Many of the Awardees and EIs we spoke with indicated that they focused their efforts on reducing PAC costs, which often involved patient education initiatives to prepare beneficiaries for a discharge home after their surgery.

The reduction in payments for orthopedic surgery episodes under Model 2 was achieved by changing the decision about where to obtain PAC. Affecting this critical decision was probably made easier by the fact that orthopedic surgery typically is elective and scheduled, allowing clinicians involved in the episode to educate the beneficiary about the best site of care. Beyond this decision, there was little evidence that EIs are affecting care after the hospital discharge. Some hospital representatives told us that they had bolstered discharge planning and case management to include periodic telephone calls with BPCI patients after their hospital discharge and to help

with making follow up appointments. However, few indicated that they actively managed the care for their patients after the PAC decision.

Medicare payments for spinal surgery episodes initiated under Model 2 went up relative to payments for episodes at comparison hospitals. At the same time, there was a statistically significant decline in mortality for BPCI episodes relative to comparison episodes. These results raise questions about why payments would go up for providers that have incentives to lower total spending. If these results persist as our sample increases, we will investigate the reasons for the increased spending and whether they are related to the decline in mortality through site visit interviews with hospital and physicians. (More recent results indicate that the reduced mortality did not persist.)

There was a statistically significant reduction in SNF length of stay for some Model 3 SNF-initiated episodes, although this did not reduce total episode payments. This type of change is directly under the SNF EI's control and would not necessarily require the EI to coordinate or manage care across the episode. Reducing SNF length of stay directly reduces SNF Medicare revenues, which could make this a difficult trade-off for SNF EIs. They would need to calculate their potential financial rewards under BPCI for reducing length of stay in comparison to foregone daily Medicare payments.

The claims-based measures of quality provide no indication of systematic effects of BPCI across most clinical episode groups. For Model 2 cardiovascular surgery and spinal surgery episodes, there were statistically significant changes in mortality relative to the comparison group, although because of baseline differences in mortality rates between BPCI and comparison hospitals, these results may be more reflective of different trends between BPCI and comparison hospitals rather than differences due to BPCI. (More recent analyses indicate that there were no differences in relative mortality rates between BPCI and comparison hospitals for these two clinical groups.) Refinements to the comparison group selection methods, which will be possible as samples increase, will allow further investigation of relative changes in mortality. Further, we will examine any corroborating information that any change is related to BPCI. Unplanned readmission rates generally have been declining, although we have noted instances in which the decline was slower for Model 3 participants than their comparison counterparts. We are not able to evaluate quality or beneficiary functioning across PAC settings. Within the separate PAC settings, there were few instances in which there were statistically significant differences in the change in functioning between BPCI and comparison patients, although when there was a statistically significant difference, the BPCI patients exhibited less improvement in functioning. Particularly with the shift of beneficiaries from institutional PAC to home health care, we will continue to examine any changes in functioning in the context of likely changes in patient complexity across PAC settings. Shifts away from institutional PAC will tend to increase patient complexity in both home and institutional PAC settings.

The beneficiary survey results indicate that BPCI has not had a negative effect on beneficiary satisfaction with their episode of care. Even for beneficiaries with MJRLE episodes, which are within the group of orthopedic surgery episodes that exhibited relative payment declines, satisfaction with care was the same between those with BPCI episodes and the comparison providers. Further, self-reported functioning improved more for beneficiaries with BPCI MJRLE

episodes in two key mobility measures. This is consistent with what many clinicians told us, that recovery from orthopedic surgery is better achieved in the beneficiary's home.

C.   Limitations

This is an observational study that relies on the differential change in claims-based and patient assessment-based measures between the BPCI participants and a comparison group to infer the impact of BPCI. The strength of these results therefore is dependent on how well the comparison group represents what would have happened absent the BPCI initiative. We have matched providers and episodes on key factors identified in the literature and by subject matter experts. We used primarily Medicare administrative data for this matching. As sample sizes increase, as well as our understanding of important differences across providers and markets, our methods for choosing a comparison group will improve. The beneficiary survey results were relative to a comparison group, although the matching process was not as rigorous as for the claims-based analysis, and was based on beneficiary recall to simulate a pre/post design. Although there are limitations with each of these designs, the strength of our mixed methods approach is the ability to triangulate results across analyses. Combined with the information gleaned from site visits, interviews, and insights from experts on the team and through expert panels, consistency across findings lend strength to our conclusions, inconsistencies raise questions for further inquiry.

It is also important to keep in mind that although this is the second Annual Report under BPCI, participants still have not had much experience under the initiative. The outcomes we report reflect up to a year under the initiative for most Awardees and EIs. The changes in care processes that are expected to change an episode of care are difficult to implement and require time to see their impact. Further, these are the earliest participants under the initiative and may not represent the large influx of participants that joined in April and July 2015.

Finally, this analytic approach is measuring multiple outcomes across a wide range of Model, participant, and episode combinations. Just by chance, some results will appear significant. This makes it even more important to compare and analyze results across outcomes, analytic approaches, and across time.

This evaluation is on-going. Additional experience under the initiative and larger sample sizes will add strength to conclusions about the impact of BPCI. Further, we are refining our analytic methods, particularly with respect to selection of comparison populations. These changes will be detailed in our next Annual Report.

D.   Future Analyses

One of the most important advances over the next year will be in our ability to examine outcomes across less aggregated clinical groups. With additional episodes and participants, we will begin to disaggregate outcomes to more meaningful clinical episodes. With a larger sample, we will also be able to expand our analyses to better understand the key characteristics of participants that contribute to their ability to achieve savings and maintain quality under BPCI. We will be able to expand our market analyses, to better understand how participation in BPCI affects participant market share and, eventually, how it affects provider referral patterns. We will also be able to examine the impact of BPCI on beneficiary subpopulations.

This Annual Report has identified several issues that require further follow up either because the finding was not expected or because it might indicate a negative consequence of the initiative. For example, we identified instances where unplanned readmissions declined at a slower pace for BPCI participants than comparison providers. We will monitor this closely to determine if this is anomalous, or whether there is a problem with certain episodes or providers under BPCI.

There are additional analyses that we will undertake when there is sufficient sample. We will examine whether the use of the 3-day hospital waiver affects the use of Medicaid-covered nursing facility care. This will require a sufficient sample of Medicare-Medicaid enrollees and more experience using the waiver prior to a SNF stay. There may be subpopulations of Medicare beneficiaries that warrant special attention under BPCI if there are any indications that their quality of care may be compromised.

E. Conclusions

This second Annual Report provides a summative and formative evaluation of the BPCI initiative based on quantitative analyses of Phase 2 participants that joined the initiative during the first year and qualitative analyses of participants that joined during the first seven quarters. Most results are based on the experience of 94 Awardees across three Models, with 227 episode initiators that were responsible for 58,410 episodes of care during the first year of the initiative. We remain limited in our ability to estimate the impact of the initiative under most Model and episode combinations because of insufficient sample size and the limited time the initiative has been underway.

Participation in BPCI has continued to grow, with more providers entering Phase 2 in April and July 2015, and more EIs transitioning episodes to Phase 2. The increased sample sizes, as well as extended times under the initiative, will allow us to expand our understanding of the impact of BPCI and strengthen our conclusions about participation under Models 2 and 3. We will use this additional sample to make more distinctions across clinical episodes and understand the experience of subpopulations that may be more vulnerable to declines in quality. Participation in Model 4, however, has dropped, which may reduce what we can say about Model 4 episodes in next year's report.

VII. References

About the National Health Interview Survey. *Centers for Disease Control and Prevention*. June 13, 2012. http://www.cdc.gov/nchs/nhis/about_nhis.htm

Brazier, JE, R Harper, NM Jones, A O'Cathain, KJ Thomas, T Usherwood, and L Westlake. "Validating the SF-36 health survey questionnaire: new outcome measure for primary care." *British Medical Journal*, Vol. 35 (6846), 160-164.

Dehejia, R.H., & Wahba, S. (2002). Propensity score-matching methods for nonexperimental causal studies. *The Review of Economics and Statistics* (84.1), 151-161.

Gage, Barbara, Roberta Constantine, Jyoti Aggarwal, Melissa Morley, vera Kurlantzick, Shulamit Bernard, Dianne Munevar, Megan Garrity, Laura Smith, Dan Bach, Anne Deutsch, Trudy Mallionson, and Linda Ehrlich-Jones. (August 2012). The Development and Testing of the Continuity Assessment Record and Evaluation (CARE) Item Set: Final Report on the Development of the CARE Item Set. *RTI International.* https://www.cms.gov/Medicare/Quality-Initiatives-Patient-Assessment-Instruments/Post-Acute-Care-Quality-Initiatives/Downloads/The-Development-and-Testing-of-the-Continuity-Assessment-Record-and-Evaluation-CARE-Item-Set-Final-Report-on-the-Development-of-the-CARE-Item-Set-Volume-1-of-3.pdf

Greene, W.H. (2002). Econometric Analysis, Fourth Edition, Prentice-Hall, Upper Saddle River, New Jersey.

Smith, J.A., & Todd, P.E. (2005). Does matching overcome LaLonde's critique of nonexperimental estimators? *Journal of Econometrics* (125), 305-353.

Tsai TC, Joynt KE, Wild RC, Oray EJ, Iha AK. Medicare's bundled payment initiative: most hospitals are focused on a few high-volume conditions. *Health Aff* 2015;34(3):3371-380. doi: 10.1377/hlthaff.2014.0900.

Zhao, Zhong. (2004). Using Matching to Estimate Treatment Effects: Data Requirements, Matching Metrics, and Monte Carlo Evidence. *Review of Economics and Statistics* (86.1), 91-107.